

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 17:28:10 ; Search time 56 Seconds

(without alignments)
1713.222 Million cell updates/sec

Title: US-09-978-385-2_COPY_19_738

Sequence: 1 STRIEQAKTFIDKFNHEAD.....DNSLEFLDIQPLGPPNPP 720

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Listing first 45 summaries

1: A.Geneseq_101002:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	3869	100.0	805 21	AA84562
2	3869	100.0	805 21	AA84562
3	3869	100.0	805 22	AA84562
4	3869	100.0	805 22	AA84562
5	3869	100.0	805 22	AA84562
6	3869	100.0	805 23	AA84562
7	3869	100.0	805 23	AA84562
8	3869	100.0	805 22	AA84562
9	3869	100.0	805 22	AA84562
10	3869	100.0	805 22	AA84562

11	2505	64.7	480	21	AA67311	Human MPR015 amin
12	1359	35.1	261	22	AA09102	Novel human protel
13	1337	34.6	732	12	AA10426	Human testicular a
14	1337	34.6	1265	23	AA020501	Protein of App rel
15	1337	34.6	1306	11	AA04111	Human angiotensin
16	1336	34.5	1306	19	AA068155	Human angiotensin
17	1334	34.5	1249	22	AA002957	Angiotensin conver
18	1334	34.5	1252	22	AA002985	Angiotensin conver
19	1069	27.6	615	22	ABB64006	Angiotensin conver
20	1054	27.2	660	16	AA070013	Drosophila melanog
21	1025.5	26.5	630	22	ABB5165	tick carboxypeptid
22	989	25.6	694	21	AA093619	Drosophila melanog
23	989	25.6	694	22	AA008701	Amino acid sequenc
24	841	21.7	792	23	ABB09432	Human zinc metallo
25	841	21.7	792	22	ABB71128	zinc metalloprotea
26	735	19.0	235	22	AA025463	Drosophila melanog
27	715.5	18.5	465	22	AA067296	Human mdt protein
28	656	17.0	252	23	AA020950	Amino acid sequenc
29	491	12.7	661	22	ABB59152	Protein of the ang
30	458	11.8	611	22	ABB70388	Drosophila melanog
31	407.5	10.5	628	22	ABB59496	Drosophila melanog
32	306	7.9	58	22	ABB38588	Peptide #6094 enco
33	306	7.9	58	22	AA059214	Human brain expres
34	306	7.9	58	22	AA071749	Human bone marrow
35	306	7.9	58	22	AA032041	Peptide #6078 enco
36	306	7.9	58	22	ABG41563	Human peptide enco
37	272	7.0	184	23	AA023565	Human EST encoded
38	272	7.0	212	21	AA033447	Human PRO1312 prot
39	272	7.0	212	21	AA034430	Human PRO1312 prot
40	272	7.0	212	21	AA06751	Membrane-bound pro
41	272	7.0	212	22	AA012412	Human PRO1312 poly
42	272	7.0	212	22	AA065274	Human PRO1312 (UNG
43	272	7.0	222	20	AA059653	Secreted protein e
44	272	7.0	222	20	AA025459	Human secreted pro
45	272	7.0	222	20	AA035890	Extended human sec

ALIGNMENTS

RESULT 1	AA84562
ID	AA84562 standard; Protein: 805 AA.
XX	AA84562;
AC	25-JUL-2000 (first entry)
XX	
DT	
XX	
DE	A human angiotensin converting enzyme-2 (ACE-2) protein.
XX	
KW	Human; angiotensin converting enzyme-2; ACE-2; angiotensin I; Ang.(1-9);
KW	blood pressure; hypertension; congestive heart failure; atherosclerosis;
KW	chronic heart failure; acute heart failure; myocardial infarction;
KW	renal failure.
XX	
OS	Homo sapiens.
XX	
FX	
FX	
FT	Key
FT	Peptide
FT	Location/Qualifiers
FT	1..18
FT	/note="signal sequence"
FT	19..740
FT	/note="extracellular domain"
FT	374..378
FT	/note="minimal zinc binding domain"
FT	741..765
FT	/note="transmembrane domain"
FT	766..805
FT	/note="cytoplasmic domain"
PN	W0200018899-A2.
XX	
PD	06-APR-2000.
XX	

PF 29-SEP-1999; 99WO-US22976.
 XX
 PR 30-SEP-1998; 98US-0163648.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Acton LS, Robison KE, Hsieh FY;
 XX
 DR WPI: 2000-293140/25.
 XX N-PSDB; AAA12764.
 XX
 PT Isolated nucleic acid encoding angiotensin converting enzyme-2 (ACE-2)
 PT polypeptide useful for detecting an ACE-2 therapeutic for treating
 PT hypertension, congestive heart failure, myocardial infarction,
 PT atherosclerosis and renal failure -
 XX
 XX Claim 2; Fig 1; 138pp: English.
 XX
 CC The present sequence represents a human angiotensin converting enzyme-2
 CC (ACE-2). ACE-2 is expressed predominantly in kidneys and testis. The
 CC sequence of the full length ACE-2 cDNA was determined from a clone
 CC obtained from a cDNA library prepared from mRNA of a human heart of
 CC a subject who had congestive heart failure. ACE-2 has significant
 CC sequence homologies with ACE enzymes, and has also been shown to
 CC hydrolyse angiotensin I into Ang.(1-9). The ACE-2 therapeutics are
 CC used to treat blood pressure related diseases and conditions, such as
 CC hypertension, congestive heart failure, chronic heart failure, acute
 CC heart failure, myocardial infarction, atherosclerosis and renal
 CC failure.
 CC
 XX
 XX Sequence 805 AA:
 SQ
 Query Match 100.0%; Score 3869; DB 21; Length 805;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEEOAKTFDKFHEADLFYQSSLASMTNTNTEENYQNNNAGDKMSAFLEKQST 60
 DB STEEOAKTFDKFHEADLFYQSSLASMTNTNTEENYQNNNAGDKMSAFLEKQST 78
 QY 19 LAQMTPLQEIOWLVKQLQALQNGSSVLSSEKSKRLNTLNTMSTIYSTGKVCNDNP 120
 DB LAQMTPLQEIOWLVKQLQALQNGSSVLSSEKSKRLNTLNTMSTIYSTGKVCNDNP 138
 QY 79 LAQMTPLQEIOWLVKQLQALQNGSSVLSSEKSKRLNTLNTMSTIYSTGKVCNDNP 138
 DB LAQMTPLQEIOWLVKQLQALQNGSSVLSSEKSKRLNTLNTMSTIYSTGKVCNDNP 138
 QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKOLRPLYEEYVVLKEMARAHYED 180
 DB 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKOLRPLYEEYVVLKEMARAHYED 198
 QY 139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKOLRPLYEEYVVLKEMARAHYED 198
 DB 139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKOLRPLYEEYVVLKEMARAHYED 198
 QY 181 YGDYWRGDEYVNGVDYDSRGQLIEDVEHTEFEIKPLYEHLHAAYRAKLMAAYPSYISP 240
 DB 181 YGDYWRGDEYVNGVDYDSRGQLIEDVEHTEFEIKPLYEHLHAAYRAKLMAAYPSYISP 258
 QY 199 YGDYWRGDEYVNGVDYDSRGQLIEDVEHTEFEIKPLYEHLHAAYRAKLMAAYPSYISP 258
 DB 199 YGDYWRGDEYVNGVDYDSRGQLIEDVEHTEFEIKPLYEHLHAAYRAKLMAAYPSYISP 258
 QY 241 IGCPLPAHLGDMGGRWTNLYSLTVPFGOKPNIDVDAMVDQAMQRIFFKEAEKFEYSV 300
 DB 241 IGCPLPAHLGDMGGRWTNLYSLTVPFGOKPNIDVDAMVDQAMQRIFFKEAEKFEYSV 318
 QY 259 IGCPLPAHLGDMGGRWTNLYSLTVPFGOKPNIDVDAMVDQAMQRIFFKEAEKFEYSV 318
 DB 259 IGCPLPAHLGDMGGRWTNLYSLTVPFGOKPNIDVDAMVDQAMQRIFFKEAEKFEYSV 318
 QY 301 GLPNMTQGFWEENSMLTDPGNVQAVCHPTAMDLGKDFRILMCTKYVMDPFLTAHHMGH 360
 DB 301 GLPNMTQGFWEENSMLTDPGNVQAVCHPTAMDLGKDFRILMCTKYVMDPFLTAHHMGH 378
 QY 319 GLPNMTQGFWEENSMLTDPGNVQAVCHPTAMDLGKDFRILMCTKYVMDPFLTAHHMGH 378
 DB 319 GLPNMTQGFWEENSMLTDPGNVQAVCHPTAMDLGKDFRILMCTKYVMDPFLTAHHMGH 378
 QY 361 IORDMAVAAQPFLLRNGANGFEHVAEGLMSTLSAAPKHLKSTGLSPDQENETELNF 420
 DB 361 IORDMAVAAQPFLLRNGANGFEHVAEGLMSTLSAAPKHLKSTGLSPDQENETELNF 438
 QY 379 IORDMAVAAQPFLLRNGANGFEHVAEGLMSTLSAAPKHLKSTGLSPDQENETELNF 438
 DB 379 IORDMAVAAQPFLLRNGANGFEHVAEGLMSTLSAAPKHLKSTGLSPDQENETELNF 438
 QY 421 LKQALITVGTLPFTVLEKRMWVFGELIPKQDMKKWEMKREITVGVPEVPHDETVC 480
 DB 421 LKQALITVGTLPFTVLEKRMWVFGELIPKQDMKKWEMKREITVGVPEVPHDETVC 498
 QY 439 LKQALITVGTLPFTVLEKRMWVFGELIPKQDMKKWEMKREITVGVPEVPHDETVC 498
 DB 439 LKQALITVGTLPFTVLEKRMWVFGELIPKQDMKKWEMKREITVGVPEVPHDETVC 498
 QY 481 DPASLFEVNSDYSEFIRYRTLYQFOQALCOAANHEGFLAKCDISNSTEAGOKLFNNL 540
 DB 481 DPASLFEVNSDYSEFIRYRTLYQFOQALCOAANHEGFLAKCDISNSTEAGOKLFNNL 558
 QY 499 DPASLFEVNSDYSEFIRYRTLYQFOQALCOAANHEGFLAKCDISNSTEAGOKLFNNL 558
 DB 499 DPASLFEVNSDYSEFIRYRTLYQFOQALCOAANHEGFLAKCDISNSTEAGOKLFNNL 558
 QY 541 RLKSEFWTLALENVGAKKMNVRPLLTNTEPELFTWLKQDNKNSFVGWSTDMSPYADQSI 600
 DB 541 RLKSEFWTLALENVGAKKMNVRPLLTNTEPELFTWLKQDNKNSFVGWSTDMSPYADQSI 600

DB 559 RLKSEFWTLALENVGAKKMNVRPLLTNTEPELFTWLKQDNKNSFVGWSTDMSPYADQSI 618
 QY 601 KVRISLSALGDKAYENDNEMTLFRSSVAYARQYELKYNOMILFGEEDYRVANILKPR 660
 DB 619 KVRISLSALGDKAYENDNEMTLFRSSVAYARQYELKYNOMILFGEEDYRVANILKPR 678
 QY 661 ISFNEFYAPKNSDIIPTREVEKAIMRSSRINDAFRLDNLSEFLGIQPTLGPNOBP 720
 DB 679 ISFNEFYAPKNSDIIPTREVEKAIMRSSRINDAFRLDNLSEFLGIQPTLGPNOBP 738

RESULT 2
 AAY67310
 ID AAY67310 standard; Protein; 805 AA.
 XX
 XX AAY67310;
 XX
 DT 11-APR-2000 (first entry)
 XX
 DE Human MPROT15 amino acid sequence #1.
 XX
 KW MPROT15; treatment; hypertension; human; myocardial disease; apoplexy;
 KW heart disease; apoplexy; heart disease; nervous denaturation; hormone;
 KW Alzheimer's disease; cytokine.
 XX
 OS Homo sapiens.
 XX
 PN JP11318472-A.
 XX
 PD 24-NOV-1999.
 XX
 PF 22-JAN-1999; 99JP-0014949.
 XX
 PR 13-MAY-1998; 98GB-0010373.
 PR 18-AUG-1998; 98GB-0018009.
 XX
 XX (SMIT) SMITHKLINE BEECHAM PLC.
 XX
 DR WPI: 2000-109268/10.
 DR N-PSDB; AA259465.
 XX
 PT MPROT15 polypeptide and MPROT15 polynucleotides - useful for the
 PT treatment of hypertension, myocardial diseases, apoplexy, heart
 PT diseases, nervous denaturation, Alzheimer's disease etc.
 XX
 PS Claim 1; Page 15; 22pp; Japanese.
 XX
 CC This is amino acid sequence #1 of human MPROT15. The MPROT15
 CC polynucleotide and polypeptide sequences can be used for the treatment of
 CC hypertension, myocardial diseases, apoplexy, heart diseases, nervous
 CC denaturation, Alzheimer's disease and diseases related to the processing
 CC of peptide hormones and cytokines.
 CC
 XX
 XX Sequence 805 AA:
 SQ
 Query Match 100.0%; Score 3869; DB 21; Length 805;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEEOAKTFDKFHEADLFYQSSLASMTNTNTEENYQNNNAGDKMSAFLEKQST 60
 DB STEEOAKTFDKFHEADLFYQSSLASMTNTNTEENYQNNNAGDKMSAFLEKQST 78
 QY 19 LAQMTPLQEIOWLVKQLQALQNGSSVLSSEKSKRLNTLNTMSTIYSTGKVCNDNP 120
 DB 19 LAQMTPLQEIOWLVKQLQALQNGSSVLSSEKSKRLNTLNTMSTIYSTGKVCNDNP 138
 QY 79 LAQMTPLQEIOWLVKQLQALQNGSSVLSSEKSKRLNTLNTMSTIYSTGKVCNDNP 138
 DB 79 LAQMTPLQEIOWLVKQLQALQNGSSVLSSEKSKRLNTLNTMSTIYSTGKVCNDNP 138
 QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKOLRPLYEEYVVLKEMARAHYED 180
 DB 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKOLRPLYEEYVVLKEMARAHYED 198
 QY 139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKOLRPLYEEYVVLKEMARAHYED 198
 DB 139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKOLRPLYEEYVVLKEMARAHYED 198
 QY 181 YGDYWRGDEYVNGVDYDSRGQLIEDVEHTEFEIKPLYEHLHAAYRAKLMAAYPSYISP 240
 DB 181 YGDYWRGDEYVNGVDYDSRGQLIEDVEHTEFEIKPLYEHLHAAYRAKLMAAYPSYISP 240

(MILL-) MILLENNIUM PHARM INC.
 Acton ST., Robison KE;
WPI: 2001-210604/21.
 DR N-PSDB; AAD02758.
 Novel genes encoding angiotensin converting enzyme-2 useful as antisense or antigene agents for therapeutics, diagnostics and screening assays -
 Claim 33; Fig 1; 76pp; English.
 The present amino acid sequence is human angiotensin converting enzyme-2 (ACE-2), also referred as peptidyl dipeptidase A (EC 3.4.15.1). Nucleic acid sequence encoding ACE-2 is useful as antisense or antigene agents for sequence specific modulation of gene expression or in the analysis of single base-pair mutations in the gene. Nucleic acid sequence encoding ACE-2 is useful in therapeutics, diagnostics and in screening assays. ACE-2 antagonist is used to treat hypertension or congestive heart failure (CHF). ACE agonist is used to reduce the inflammation and pain resulting from an insect sting or bite, which was accompanied by an injection of bradykinin. Anti-ACE-2 antibodies are used to monitor ACE-2 protein levels for determining the disease or condition associated with an aberrant protein level.

Sequence	805 AA;
Query Match	100.0%; Score 3869; DB 22; Length 805;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 720; Conservative	0; Mismatches 0; Indels 0; Gaps 0;

Db	19	STIEQATFELDKRNEAEEDLFYQSSIASMNNTNTEEVQNMNNAAGDKMSAFLLKQST	78
QY	61	LAQMYPLAQEIQNLVVKLTOLALQONSSVLSDEKSKRLNTLINTMSTYISYGGKACPNDP	120
Db	79	LAQMYPLAQEIQNLVVKLTOLALQONSSVLSDEKSKRLNTLINTMSTYISYGGKACPNDP	138
QY	121	QECILLEPGLEIMANSLDYNERLWAMESRSEVQKQRLPYEEYVLKNEMAANHYED	180
Db	139	QECILLEPGLEIMANSLDYNERLWAMESRSEVQKQRLPYEEYVLKNEMAANHYED	198
QY	181	YGDYWRDGYEVNVDGIDYSRGOLLIEDVHTTEETKPLYEHLHAAYRAKLMAAPYSISP	240
Db	199	YGDYWRDGYEVNVDGIDYSRGOLLIEDVHTTEETKPLYEHLHAAYRAKLMAAPYSISP	258
QY	241	IGCLPAHLIGDMGRFMTNLISLTPFGQKPNIDVTDAWQANDADRIKREAEKFPVSY	300
Db	259	IGCLPAHLIGDMGRFMTNLISLTPFGQKPNIDVTDAWQANDADRIKREAEKFPVSY	318
QY	301	GLPNNQGFENSNMLDPGCVQKAYCHPPLANDLKGDFILMCTKVTIMDDFLTAHHNMG	360
Db	319	GLPNNQGFENSNMLDPGCVQKAYCHPPLANDLKGDFILMCTKVTIMDDFLTAHHNMG	378
QY	361	IQYDMAVYAAQPEFLLRKNGANEHVEYAGTMSLSATPRHLKSTIGLLSPDQOEINTEFIN	420
Db	379	IQYDMAVYAAQPEFLLRKNGANEHVEYAGTMSLSATPRHLKSTIGLLSPDQOEINTEFIN	438
QY	421	LILQALTYIGTLPTTYMLERKMWVFKGELPRDOMKRWKEMKREITGVGVEPRPHETVC	480
Db	439	LILQALTYIGTLPTTYMLERKMWVFKGELPRDOMKRWKEMKREITGVGVEPRPHETVC	498

481	DPASLEPHVNSDYSFTRYTRRLYPOFOPEALCOAAKHGEPGLHCDISNSTEAGOKLENNL	540
482		
483		
484		
485		
486		
487		
488		
489	DPASLEPHVNSDYSFTRYTRRLYPOFOPEALCOAAKHGEPGLHCDISNSTEAGOKLENNL	558
490		
491		
492		
493		
494		
495		
496		
497		
498		
499		
500		
501		
502		
503		
504		
505		
506		
507		
508		
509		
510		
511		
512		
513		
514		
515		
516		
517		
518		
519		
520		
521		
522		
523		
524		
525		
526		
527		
528		
529		
530		
531		
532		
533		
534		
535		
536		
537		
538		
539		
540		
541	RCKSEPPTALLENVVGAKNNNVPLNYPPEPLTKDKDNKNSFVGWSTIDMSPYADDSI	600
542		
543		
544		
545		
546		
547		
548		
549		
550		
551		
552		
553		
554		
555		
556		
557		
558		
559		
560		
561		
562		
563		
564		
565		
566		
567		
568		
569		
570		
571		
572		
573		
574		
575		
576		
577		
578		
579		
580		
581		
582		
583		
584		
585		
586		
587		
588		
589		
590		
591		
592		
593		
594		
595		
596		
597		
598		
599		
600		

QY 601 KVRISLKSALGDKAYENNENMTLFRSSVAYAMROYFLKKNOMILFGEEDVAVANLKPR 660
 DB 619 KVRISLKSALGDKAYENNENMTLFRSSVAYAMROYFLKKNOMILFGEEDVAVANLKPR 678
 QY 661 ISFNEFVYAPKNVSDIIPRTEVEKAIMRSRINDAFLRNDLSLEFLGIOTPLGPPNQP 720
 DB 679 ISFNEFVYAPKNVSDIIPRTEVEKAIMRSRINDAFLRNDLSLEFLGIOTPLGPPNQP 738

RESULT 4
 AAB48095 standard; Protein: 805 AA.
 AAB48095;
 AAB48095;
 19-MAR-2001 (first entry)
 Human Zace2 protein.
 Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
 zinc metalloproteinase; blood pressure; zinc protease; hypertension;
 ventricular systolic dysfunction; renal impairment; heart failure;
 scleroderma renal crisis; atherosclerosis; antiinflammatory; human;
 antiarthritis; bradykinin inactivator.
 Homo sapiens.
 WO200070032-A1.
 23-NOV-2000.
 03-MAY-2000; 2000MO-US11932.
 13-MAY-1999; 9905-0311482
 27-AUG-1999; 9905-0384706
 (ZYMO) ZYMOGENETICS INC.
 Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
 WPI: 2001-025018/03.
 N-PSDB; AAC84366, AAC84367.
 Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 associated with inflammation such as arthritis and enterocolitis -
 Example 1; Page 95-100; 125pp; English.
 The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-
 converting enzyme is a zinc metalloproteinase that plays roles in blood
 pressure regulation and fertility. Zace2 can be expressed by standard
 recombinant methodology. Zace2 polypeptides are useful for treating an
 inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
 as targets for identifying modulators of zinc protease activity, for
 screening or identifying new angiotensin-converting enzyme (ACE)
 inhibitors, and as a basis for rational drug design for inhibitory
 molecules. The nucleic acids can be used to detect the expression of a
 Zace2 gene in a biological sample, as probes for in vivo diagnosis and
 for detecting and localizing Zace2 gene expression in tissue samples,
 to determine whether a subject's chromosomes contain a mutation in the
 Zace2 gene, and to detect aberrations associated with the Zace2 locus.
 Inhibitors of ACE are used for treating hypertension of various
 conditions, including left ventricular systolic dysfunction, progressive
 renal impairment, scleroderma renal crisis, congestive heart failure due
 to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
 used to treat infertility while Zace2 antagonists are used for inducing
 infertility. The present sequence represents the human Zace2 protein.

Sequence 805 AA;
 Query Match 100.0%; Score 3869; DB 22; Length 805;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STIEBOAKTFLDKFHEAEADLFYQSSLASWNTNTTEENYQNNMAGDKASATLQST 60
 DB 19 STIEBOAKTFLDKFHEAEADLFYQSSLASWNTNTTEENYQNNMAGDKASATLQST 78
 QY 61 LKOWPBLQETQWLYKLOLQALQONGSSVLSBDSKRLNTLNTSTYTGKVCNPNP 120
 DB 79 LKOWPBLQETQWLYKLOLQALQONGSSVLSBDSKRLNTLNTSTYTGKVCNPNP 138
 QY 121 QECLELPEGLNEMANSIDYNERLWMSWRSVEGKOLRPLYEEVYVAKNMAVSYISP 240
 DB 139 QECLELPEGLNEMANSIDYNERLWMSWRSVEGKOLRPLYEEVYVAKNMAVSYISP 198
 QY 181 YGDYWRGDEYVNGVYQSDYSGQLLEDVHTFEELKPLYEHHAAYRAKIMNAPSYISP 240
 DB 199 YGDYWRGDEYVNGVYQSDYSGQLLEDVHTFEELKPLYEHHAAYRAKIMNAPSYISP 258
 QY 241 IGLPAHLIGDMNGREFTNLYSLVFPFGKPNIDVTDAMVDQAMDAQRIEAEKFEVSV 300
 DB 259 IGLPAHLIGDMNGREFTNLYSLVFPFGKPNIDVTDAMVDQAMDAQRIEAEKFEVSV 318
 QY 301 GLPMTQGEWNSMLDPCGVQAVCHPTAMDLGKDFELMCTVYTMDFLTAHENGH 360
 DB 319 GLPMTQGEWNSMLDPCGVQAVCHPTAMDLGKDFELMCTVYTMDFLTAHENGH 378
 QY 361 IYDMMAYAAQFLIRNGANGFEHAEVGEIMSLSAATPKHLKSLIGLSPDQEDNETINF 420
 DB 379 IYDMMAYAAQFLIRNGANGFEHAEVGEIMSLSAATPKHLKSLIGLSPDQEDNETINF 438
 QY 421 LKQALITVGTLPFTYMLEKRWAVKGEIPKQNMKKWEMKREIVGVVEVPDHEVC 480
 DB 439 LKQALITVGTLPFTYMLEKRWAVKGEIPKQNMKKWEMKREIVGVVEVPDHEVC 498
 QY 481 DPASLPHVNSDYSPIRYRTFLYQFOEALCOAAHHEGLHRCDISNSTEAGOKLFNML 540
 DB 499 DPASLPHVNSDYSPIRYRTFLYQFOEALCOAAHHEGLHRCDISNSTEAGOKLFNML 558
 QY 541 RLGRSEPTLALENVGAKMNTNPLNTEPEFTWLKQNKSPFGMSSTWSPYADQST 600
 DB 559 RLGRSEPTLALENVGAKMNTNPLNTEPEFTWLKQNKSPFGMSSTWSPYADQST 618
 QY 601 KVRISLKSALGDKAYENNENMTLFRSSVAYAMROYFLKKNOMILFGEEDVAVANLKPR 660
 DB 619 KVRISLKSALGDKAYENNENMTLFRSSVAYAMROYFLKKNOMILFGEEDVAVANLKPR 678
 QY 661 ISFNEFVYAPKNVSDIIPRTEVEKAIMRSRINDAFLRNDLSLEFLGIOTPLGPPNQP 720
 DB 679 ISFNEFVYAPKNVSDIIPRTEVEKAIMRSRINDAFLRNDLSLEFLGIOTPLGPPNQP 738

RESULT 5
 AAU99701 standard; Protein: 805 AA.
 AAU99701;
 24-SEP-2002 (first entry)
 Human angiotensin converting enzyme-2 (ACE-2) protein.
 Human; angiotensin converting enzyme-2; ACE-2; body weight disorder;
 muscle mass; body fat; obesity; diabetes; atherosclerosis; weight loss;
 lipid metabolism; weight gain; anorexia; cachexia; bulimia; sepsis;
 familial partial lipodystrophy; hypercholesterolemia; hyperlipidemia;
 aberrant metabolic rate; heart failure; left ventricular hypertrophy;
 neurodegenerative disorder; peptide hormone; cytokine processing;
 myocardial infarction; cardiomyopathy; inflammatory bowel disease;
 systemic inflammation response syndrome; polytrauma; pain; stroke;
 bone destruction; rheumatoid arthritis; osteoarthritis; asthma;
 periodontal disease; dysmenorrhea; premature labour; brain oedema;
 focal injury; diffuse axonal injury; reperfusion injury; scar formation;

KW cerebral vasospasm; subarachnoid haemorrhage; allergic disorder;
 KW adult respiratory distress syndrome; wound healing; appetite;
 KW body mass index.
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT /label= signal_peptide
 FT 1..18
 FT 19..805
 FT /label= Mature_human_ACE_2_protein
 PN W020039997-A2.
 PD 23-MAY-2002.
 XX
 PE 31-OCT-2001; 2001MO-US45703.
 XX
 PR 01-NOV-2000; 2000US-0704216
 PR 29-MAY-2001; 2001US-087038Z.
 PR 19-OCT-2001; 2001US-371741P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 PI Acton ST, Ocala TD, Gould AE, Dales NA, Guan B, Brown JA;
 PI Petane M, Kadambi VJ, Solomon M, Stricker-Krongrad A;
 XX
 DR WPI; 2002-547572/58.
 DR N-PSDB; ABR87623.
 XX
 PT Treating body weight disorder and increasing muscle mass comprises
 PT administering angiotensin converting enzyme-2 modulating compound -
 XX
 PS Example 5; Page 387-390; 395pp; English.
 XX
 CC The present invention describes a new method of treating a body weight
 CC disorder, increasing muscle mass and decreasing body fat by
 CC administration of angiotensin converting enzyme (ACE)-2 modulating
 CC compound. The invention can be used for treating body weight disorders,
 CC particularly obesity of at least grade 1, diabetes, atherosclerosis and
 CC a state associated with lipid metabolism. The method is used for treating
 CC rapid weight loss, rapid weight gain, anorexia, cachexia, bulimia,
 CC generalised partial lipodystrophy, familial partial lipodystrophy,
 CC hypercholesterolaemia, hyperlipidaemia, an aberrant metabolic rate,
 CC congestive heart failure, chronic heart failure, left ventricular
 CC hypertrophy, acute heart failure, neurodegenerative disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease and Huntington's disease),
 CC diseases associated with peptide hormones or cytokine processing,
 CC myocardial infarction, cardiomyopathy, systemic inflammation response
 CC syndrome, sepsis, polytrauma, inflammatory bowel disease, acute and
 CC chronic pain, bone destruction in rheumatoid arthritis and osteoarthritis
 CC and periodontal disease, dysmenorrhoea, premature labour, brain oedema
 CC following focal injury, diffuse axonal injury, stroke, reperfusion
 CC injury, cerebral vasospasm after subarachnoid haemorrhage, allergic
 CC disorders including asthma, adult respiratory distress syndrome, wound
 CC healing and scar formation. The invention decreases the appetite,
 CC increases muscle mass and decreases body fat of subject having body mass
 CC index of greater than 23 (preferably 24.9)kg/m². The present amino
 CC acid sequence represents the human ACE-2 protein of the invention.
 XX
 SQ Sequence 805 AA:
 Query Match 100.0%; Score 3869; DB 23; Length 805;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STEEQAFTFLDKFNHEADLFYQSSIASMNTNTNTEENYONNANNAGKMSAFLEKEST 60
 DB 19 STEEQAFTFLDKFNHEADLFYQSSIASMNTNTNTEENYONNANNAGKMSAFLEKEST 78
 QY 61 LAQMTYLAQETONLYKALQALQONGSVLSSEKSKRLTILNTSTIYSTGKVCNPDP 120
 DB 79 LAQMTYLAQETONLYKALQALQONGSVLSSEKSKRLTILNTSTIYSTGKVCNPDP 138

QY 121 QECILLEPELTNEIMANSIDYNERIMAWSEWSEVKGOLRPLYEEYVVLKNEMARAHYED 180
 DB 139 QECILLEPELTNEIMANSIDYNERIMAWSEWSEVKGOLRPLYEEYVVLKNEMARAHYED 198
 QY 181 YGDYWRGDEYVNGVDYDYSRGQLIEDVEHFEETKPLYLEHIAVRAKIMNAYPSYSP 240
 DB 199 YGDYWRGDEYVNGVDYDYSRGQLIEDVEHFEETKPLYLEHIAVRAKIMNAYPSYSP 258
 QY 241 IGCLEPAHLIGDMWGFNINLSLYTPFGQKNIDVTAMDQAMDQRIKFEAEKEFVS 300
 DB 259 IGCLEPAHLIGDMWGFNINLSLYTPFGQKNIDVTAMDQAMDQRIKFEAEKEFVS 318
 QY 301 GLPNMTOGFENSMITDPCNVOKAVCHPTAMDGLGKDRITLMCTKVTDMDLLTAHHENG 360
 DB 319 GLPNMTOGFENSMITDPCNVOKAVCHPTAMDGLGKDRITLMCTKVTDMDLLTAHHENG 378
 QY 361 IOYDMAVYAAQPLLRNGANEHGEHVAEIMTSIAATPKHLKSIGLLSPQEDNETEINF 420
 DB 379 IOYDMAVYAAQPLLRNGANEHGEHVAEIMTSIAATPKHLKSIGLLSPQEDNETEINF 438
 QY 421 LKQALITVCTLPFTYMLEKRWKMFKEIIPDQMKMKWEMKREIVGVPEVPHDETYC 480
 DB 439 LKQALITVCTLPFTYMLEKRWKMFKEIIPDQMKMKWEMKREIVGVPEVPHDETYC 498
 QY 481 DPASLEFHSNDSEFRTYTRTLTYPOFOALCOQAKHEGPHKCDISNSTAGOKLFNML 540
 DB 499 DPASLEFHSNDSEFRTYTRTLTYPOFOALCOQAKHEGPHKCDISNSTAGOKLFNML 558
 QY 541 RLGKSEPTLALENVYGAKNMVRPLNYFEPLFTWLDQKNSEFVGMSTDMSPYADQSI 600
 DB 559 RLGKSEPTLALENVYGAKNMVRPLNYFEPLFTWLDQKNSEFVGMSTDMSPYADQSI 618
 QY 601 KVISLKSALGDKAEYNDNEMKYLEPSSVAARQYFLKAVNOMILBEEBYRANAKPR 660
 DB 619 KVISLKSALGDKAEYNDNEMKYLEPSSVAARQYFLKAVNOMILBEEBYRANAKPR 678
 QY 661 ISFNEFYAPKKNVSDIIRTEVEKAIKMSRSKINDAFRLNDSLEFLGIDPTLPPNQP 720
 DB 679 ISFNEFYAPKKNVSDIIRTEVEKAIKMSRSKINDAFRLNDSLEFLGIDPTLPPNQP 738
 RESULT 6
 AAE20353
 ID AAE20353 standard; Protein; 805 AA.
 XX
 AC AAE20353;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 XX Human ACE-2 full-length protein.
 DE Human: angiotensin converting enzyme-2; ACE-2; therapy; hypertension;
 KW peptidyl dipeptidase A; blood pressure; hypotension; atherosclerosis;
 KW myocardial infarction; heart failure; arrhythmia; renal failure; gene;
 KW inflammation; fertility; enzyme; EC 3.4.15.1.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT /label= signal_peptide
 FT 1..18
 FT 19..805
 FT /note= "Mature ACE-2 protein"
 FT 19..740
 FT Domain
 FT /note= "Extracellular domain"
 FT 374..378
 FT /note= "Zinc binding domain (ZBD)"
 FT 741..765
 FT /note= "Transmembrane domain"
 FT 766..805
 FT /note= "Cytoplasmic domain"
 FT Domain

PN W0200212471-A2.
 XX 14-FEB-2002.
 XX 09-AUG-2001; 2001NO-US25059.
 XX 09-AUG-2000; 2000US-0635501.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA Action S, Robison KE, Hsieh FY;
 XX WPI; 2002-257481/30.
 DR N-PSDB; AAD32586.
 XX
 PT Isolated human polypeptide, known as angiotensin converting enzyme-2,
 PT useful for treating or preventing the development of an abnormal blood
 PT pressure or related diseases, e.g. hypertension, heart failure or
 PT myocardial infarction.
 PS
 PS Claim 2; Fig 1; 218pp; English.
 CC The invention relates to human angiotensin converting enzyme-2 (ACE-2)
 CC polypeptides and polynucleotides. ACE-2 is also known as peptidyl
 CC dipeptidase A (EC 3.4.15.1). Polypeptides of the invention are useful
 CC for treating or preventing the development of abnormal blood pressure
 CC and diseases or disorders associated with the protein in a subject. The
 CC diseases include hypertension, hypotension, congestive heart failure,
 CC chronic heart failure, acute heart failure, myocardial infarction,
 CC atherosclerosis, arrhythmia and renal failure. They are also useful
 CC for treating inflammatory conditions and diseases relating to fertility.
 CC The present sequence is human full-length ACE-2 protein.
 XX
 XX Sequence 805 AA:
 Query Match 100.0%; Score 3869; DB 23; Length 805;
 Best Local Similarity 100.0%; Pred. NO. 0;
 Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STEEQAKTFLDKFNHEAEDLFYQSSLASNNYNTNITEENVQNNMAGDKMSAFLEKQST 60
 DB 19 STEEQAKTFLDKFNHEAEDLFYQSSLASNNYNTNITEENVQNNMAGDKMSAFLEKQST 78
 QY 61 LAQYPIQEIQLNFKVQLQALQONGSSVSEDKSKLNTLNTMSTTSGKVCNDNP 120
 DB 79 LAQYPIQEIQLNFKVQLQALQONGSSVSEDKSKLNTLNTMSTTSGKVCNDNP 138
 QY 121 QECLEPGLNEIMANSIDYNERLWMSRSEVKQRLPYEEYVVLKKNEMARANYED 180
 DB 139 QECLEPGLNEIMANSIDYNERLWMSRSEVKQRLPYEEYVVLKKNEMARANYED 198
 QY 181 YGDYWRGDEYNGVDYDSRGQLIEDVEHTFEEIKPLYEHLAYYRAKIMNAVSYISP 240
 DB 199 YGDYWRGDEYNGVDYDSRGQLIEDVEHTFEEIKPLYEHLAYYRAKIMNAVSYISP 258
 QY 241 IGCPRHLIDGMGRFTNLTSLTVPPGQKRNIDVTAMDOADARIFREKAEKFPVS 300
 DB 259 IGCPRHLIDGMGRFTNLTSLTVPPGQKRNIDVTAMDOADARIFREKAEKFPVS 318
 QY 301 GLPNMTOGFEMSNLTDPGNVOKAVCHPTANDLKGDFRILMCKVKYVMDDFLTAAHMGH 360
 DB 319 GLPNMTOGFEMSNLTDPGNVOKAVCHPTANDLKGDFRILMCKVKYVMDDFLTAAHMGH 378
 QY 361 IGYDMAVAAOPFLLRNGANGEGFHEAVGKLSAATPKNKLISGLLSPDEQEDNETINF 420
 DB 379 IGYDMAVAAOPFLLRNGANGEGFHEAVGKLSAATPKNKLISGLLSPDEQEDNETINF 438
 QY 421 LLKQALITVGLPFLPYMLEKRWAVFGETPKDQMMKMKREIYGVAVPVHDEYTC 480
 DB 439 LLKQALITVGLPFLPYMLEKRWAVFGETPKDQMMKMKREIYGVAVPVHDEYTC 498
 QY 481 DPASLFHNSDYSPFIRYTRFLYQFOFQALCOAKHKGPIAKHDISNSTEAGOKLPMNL 540

DB 499 DPASLFHNSDYSPFIRYTRFLYQFOFQALCOAKHKGPIAKHDISNSTEAGOKLPMNL 558
 QY 541 RLKSEPTWLALENVYGAKNMNVRLNYFEPLFTWLDQKNSEFVGMSTWSPYADQSI 600
 DB 559 RLKSEPTWLALENVYGAKNMNVRLNYFEPLFTWLDQKNSEFVGMSTWSPYADQSI 618
 QY 601 KYRISLKSALGKAVKENDNEMYLPRSSVAYAMROYELKVNQNTLFEEDYRVANLKP 660
 DB 619 KYRISLKSALGKAVKENDNEMYLPRSSVAYAMROYELKVNQNTLFEEDYRVANLKP 678
 QY 661 ISFNFPYAPKAVNSDIIPTEVEKARMSRSRINAFRLNDSLEFLGIQPTLGPNNOP 720
 DB 678 ISFNFPYAPKAVNSDIIPTEVEKARMSRSRINAFRLNDSLEFLGIQPTLGPNNOP 738
 RESULT 7
 ID AAU09092 standard; Protein; 711 AA.
 XX AAU09092;
 AC 20-DEC-2001 (first entry)
 DE Novel human protein NHP #1.
 XX Human; novel human protein; NHP; antidiabetic; antineumatic;
 XX antiallergic; cytosolic; antiarteriosclerotic; vlnarary;
 XX neuroprotective; nootropic; antiparkinsonian;
 XX anti-human immunodeficiency virus; antiaesthetic; vasotropic; cardiac;
 XX hypotensive; anorectic; antinfertility; neuroleptic; anticonvulsant;
 XX antihuman; immunosuppressive; cerebroprotective; antitubercial;
 XX antinflamatory; antibacterial; antiporiatic; thyromimetic;
 XX immunomodulator; antiseborrheic; dermatological; vasoconstriction;
 XX gastrointestinal disorder; cardiovascular disorder; hypertension;
 XX coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;
 XX cachexia; male infertility; impotence; testicular cancer; lung tumour;
 XX hyperproliferative disorder; pulmonary system disorder;
 XX central nervous system disorder; bone disorder;
 XX neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 XX Huntington's disease; schizophrenia; mania; dementia; paraneola;
 XX panic disorder; learning disability; amyotrophic lateral sclerosis;
 XX psychosis; autism; sleep disorder; immune system disorder;
 XX Hashimoto's thyroiditis; musculo-skeletal system disorders;
 XX multiple sclerosis; ischemic brain injury; stroke; infectious disease;
 XX diabetes mellitus; immunological disorder; asthma; AIDS; immunogen;
 XX acquired immunodeficient syndrome; leukaemia; rheumatoid arthritis;
 XX inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;
 XX neural system disorder; respiratory disorder; olfactory disorder;
 XX wound healing; chromosome X.
 OS Homo sapiens.
 XX
 FH Key
 FT Domain
 FT 1..681
 FT /label= Extracellular_domain
 FT 48..55
 FT /label= Immunogenic-epitope
 FT 110..118
 FT /label= Immunogenic-epitope
 FT 136..146
 FT /label= Immunogenic-epitope
 FT 151..158
 FT /label= Immunogenic-epitope
 FT 219
 FT /label= OTHER
 FT /note= "Other" Any amino acid encoded by WST"
 FT 240
 FT /label= OTHER
 FT /note= "Other" Any amino acid encoded by RGC"
 FT 499
 FT /label= OTHER
 FT /note= "Other" Any amino acid encoded by NTT"
 FT 682..698
 FT Domain

/label= Transmembrane_domain

XX FT
XX PN
XX WO200174896-A1.
XX PD
XX 11-OCT-2001.
XX PF
XX 02-APR-2001; 2001WO-US10542.
XX PR
XX 03-APR-2000; 2000US-194118P.
XX PR
XX 29-SEP-2000; 2000US-236384P.
XX PA
XX (HUMA-) HUMAN GENOME SCI INC.
XX PI
XX Moore PA, Ni J, Soppet DR, Coleman TA, Gentz RL, Endress GA;
XX Li Y, Dillon PJ;
XX WPI: 2001-626394/72.
XX N-PSDB; AAS14880.

PT New human proteins, useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cardiovascular
PT disorders, autoimmune disorders and reproductive disorders

PS Claim 11; Page 298-301; 318pp; English.

CC The invention relates to novel human proteins (NHP) and the
CC nucleic acids that encode them and antibodies raised against them.
CC The proteins, antibodies and nucleic acids are useful in the diagnosis,
CC prognosis, prevention and/or treatment of diseases and/or disorders
CC involving vasoconstriction, gastrointestinal disorders, cardiovascular
CC disorders (e.g. hypertension, erectile dysfunction, high blood pressure,
CC coronary heart disease and arteriosclerosis), anorexia, obesity, bulimia,
CC cachexia, disorders of small intestine, disorders of reproductive system
CC (e.g. male infertility and/or impotence), testicular cancer, lung tumours
CC and other hyperproliferative disorders, disorders of pulmonary system
CC central nervous system disorders, bone disorders, neurodegenerative
CC diseases and behavioural disorders (e.g. Alzheimer's disease, Parkinson's
CC disease, Huntington's disease, schizophrenia, mania, dementia, paranoia,
CC panic disorder, learning disabilities, amyotrophic lateral sclerosis,
CC psychoses, autism, sleep disorders), immune system disorders (e.g.
CC Hashimoto's thyroiditis), renal and musculo-skeletal system disorders,
CC central nervous system disorders (e.g. multiple sclerosis, ischemic
CC brain injury and/or stroke), infectious diseases, diabetes mellitus,
CC immunological disorders (e.g. asthma, acquired immunodeficient syndrome
CC (AIDS), leukaemia, rheumatoid arthritis, inflammatory bowel disease,
CC sepsis, acne, psoriasis and lupus erythematosus), neural system
CC disorders, respiratory disorders, olfactory disorders and wound
CC healing. The present sequence represents an NHP of the invention the
CC gene for which is located on the X chromosome.

XX Sequence 711 AA.

Query Match 93.6%; Score 3620; DB 22; Length 711;

Best Local Similarity 99.3%; Pred. No. 5.3e-312;

Matches 672; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

DB 44 MNMGDKMSAPLKKESTIAQMPPIQLEIONLVKVLQALQONGSSVISEDKSKRLITLN 103
1 MNMGDKMSAPLKKESTIAQMPPIQLEIONLVKVLQALQONGSSVISEDKSKRLITLN 60
QY 104 TMSITSTGKVCNPDNPQECLELLEPGLEIMANSIDYERIMAMESWSEYKOLRPIYE 163
DB 61 TMSITSTGKVCNPDNPQECLELLEPGLEIMANSIDYERIMAMESWSEYKOLRPIYE 120
QY 164 EYVLAKNEMARAHYEDYGDYWRDYEYNGVDYDYSRGQLEIDYEHTFEERIKPLYETH 223
DB 121 EYVLAKNEMARAHYEDYGDYWRDYEYNGVDYDYSRGQLEIDYEHTFEERIKPLYETH 180
QY 224 AYVRAKLNAYPSYISPGCLPAHLGDMGRFTNTLSLVPPGQRPNDIVTAMPDQA 283
DB 181 AYVRAKLNAYPSYISPGCLPAHLGDMGRFTNTLSLVPPGQRPNDIVTAMPDQA 240
QY 284 WDAORIFKEAEKFEVSVGLPMNTOGFWMENSLTDPGYNQAVCHPTAMDGLKGDFRILMC 343

DB 241 WDAORIFKEAEKFEVSVGLPMNTOGFWMENSLTDPGYNQAVCHPTAMDGLKGDFRILMC 300
QY 344 TKVTMDDFLFAHHEMGHIQDMAYAAOPFLIRNANGCEHAEIGELISAATPKHKST 403
DB 301 TKVTMDDFLFAHHEMGHIQDMAYAAOPFLIRNANGCEHAEIGELISAATPKHKST 360
QY 404 GLISPDFOEDNETEINFLKQALITVGLTFTVYLEKRMNMYKGEIPKDONMKWEMK 463
DB 361 GLISPDFOEDNETEINFLKQALITVGLTFTVYLEKRMNMYKGEIPKDONMKWEMK 420
QY 464 RELVGYEPPPHDTEYCDPASLFEVSDYSFIRYRTTLVQFOFOALCOAAHHEGPLK 523
DB 421 RELVGYEPPPHDTEYCDPASLFEVSDYSFIRYRTTLVQFOFOALCOAAHHEGPLK 480
QY 524 CDISNSTEAGOKLEFMRLKGESEPTALENVGAKMANRYPLINTEPEPLTLKQONK 583
DB 481 CDISNSTEAGOKLEFMRLKGESEPTALENVGAKMANRYPLINTEPEPLTLKQONK 540
QY 584 SFVGMSTDMSPYADQSIKVRISLKSALGDRAYENDNEMYLFRSSVAYAMROYFLVKYK 643
DB 541 SFVGMSTDMSPYADQSIKVRISLKSALGDRAYENDNEMYLFRSSVAYAMROYFLVKYK 600
QY 644 MILEGEEDYRANIKPRISNFFVTAKNVDIIPTEVEKAIKMSKRSINDAFRLNDS 703
DB 601 MILEGEEDYRANIKPRISNFFVTAKNVDIIPTEVEKAIKMSKRSINDAFRLNDS 660
QY 704 LEFGIOPTLGPPNQPP 720
DB 661 LEFGIOPTLGPPNQPP 677

RESULT 8
AAB48097
ID AAB48097 standard; Protein; 805 AA.
XX AC
XX AAB48097;
DT 19-MAR-2001 (first entry)
XX DE
XX Mouse Zace2-5 protein.
KW Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
KW zinc metalloproteinase; blood pressure; zinc protease; hypertension;
KW ventricular systolic dysfunction; renal impairment; heart failure;
KW scleroderma renal crisis; atherosclerosis; antiinflammatory; mouse;
KW antiarthritic; bradykinin inactivator.
XX OS
XX Mus sp.
XX FH
XX Key
FH 19..613
FT Region
FT /note= "fragment specifically claimed for"
FT 19..708
FT Region
FT /note= "fragment specifically claimed for"
FT 19..738
FT Region
FT /note= "fragment specifically claimed for"
FT 19..805
FT Region
FT /note= "fragment specifically claimed for"
FT 133..542
FT Region
FT /note= "fragment specifically claimed for"
FT 344..542
FT Region
FT /note= "fragment specifically claimed for"
FT 371..402
FT Region
FT /note= "fragment specifically claimed for"
PN WO200070032-A1.
XX 23-NOV-2000.
PD 03-MAY-2000; 2000WO-US11932.
PF 13-MAY-1999; 99US-0311482.
PR

PR 27-AUG-1999; 99US-0384706.
 XX (ZYMO) ZYMOGENETICS INC.
 PA Piddington CS, Petrie CR, Shoemaker KE, Bishop PD.
 PI WPI; 2001-025018/03.
 DR N-PSDB; AAC84368, AAC84369.
 XX
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 PT associated with inflammation such as arthritis and enterocolitis -
 XX
 PS Claim 7; Page 104-109; 125pp; English.
 XX
 CC The invention relates to the metalloenzyme zace2, zace2, an angiotensin-
 CC converting enzyme is a zinc metalloprotease that plays roles in blood
 CC pressure regulation and fertility. Zace2 can be expressed by standard
 CC recombinant methodology. Zace2 polypeptides are useful for treating an
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
 CC diseases associated with inflammation like arthritis and enterocolitis,
 CC as targets for identifying modulators of zinc protease activity, for
 CC screening or identifying new angiotensin-converting enzyme (ACE)
 CC inhibitors, and as a basis for rational drug design for inhibitory
 CC molecules. The nucleic acids can be used to detect the expression of a
 CC zace2 gene in a biological sample, as probes for in vivo diagnosis and
 CC for detecting and localizing zace2 gene expression in tissue samples,
 CC to determine whether a subject's chromosomes contain a mutation in the
 CC zace2 gene, and to detect aberrations associated with the zace2 locus.
 CC Inhibitors of ACE are used for treating hypertension of various
 CC conditions, including left ventricular systolic dysfunction, progressive
 CC renal impairment, scleroderma renal crisis, congestive heart failure due
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
 CC used to treat infertility while zace2 antagonists are used for inducing
 CC infertility. The present sequence represents the mouse zace2-5 protein.
 CC
 XX
 SQ Sequence 805 AA;
 Query Match 84.0%; Score 3251; DB 22; Length 805;
 Best Local Similarity 83.1%; Pred. No. 3.6e-279;
 Matches 598; Conservative 49; Mismatches 73; Indels 0; Gaps 0;

QY 481 DPASLEHNSDYSEFRTYRTLYQFOFQALCOAAKHGPHAKCDISNSTEAGOKLPM 540
 DB 499 DPASLEHNSDYSEFRTYRTLYQFOFQALCOAAKHGPHAKCDISNSTEAGOKLPM 558
 QY 541 RLGKSEPTWALLENVGAANNVRLPLNFEPLFTWLDKQNSFVQSTWSPYADQSI 600
 DB 559 SLGNSEPTWALLENVGAANNVRLPLNFEPLFTWLDKQNSFVQSTWSPYADQSI 618
 QY 601 KVRISLSKSLGDAKAVEMNDNEMYLFRSSVAYAMROYFLKVNOMLIGGEEDVRANLKR 660
 DB 619 KVRISLSKSLGDAKAVEMNDNEMYLFRSSVAYAMROYFLKVNOMLIGGEEDVRANLKR 678
 QY 661 ISFNEFVAPKKNVSDIIPTEVEKARMSRSRINDAFRLNDSLEFLIOPTLGPNNQP 720
 DB 679 VSFYFVFVSPQVSDVIRSEVEDAIRMSRGRINDVFLNDSLEFLIOPTLGPNNQP 738
 RESULT 9
 AAB48098
 ID AAB48098 standard; Protein; 805 AA.
 XX
 AC AAB48098;
 XX
 DT 19-MAR-2001 (first entry)
 XX
 DE Mouse Zace2-10 protein.
 XX
 KW Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
 KW zinc metalloprotease; blood pressure; zinc protease; hypertension;
 KW ventricular systolic dysfunction; renal impairment; heart failure;
 KW scleroderma renal crisis; atherosclerosis; antiinflammatory; mouse;
 KW antihypertic; bradykinin inactivator.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT Region 19..613
 FT /note= "fragment specifically claimed for"
 FT Region 19..708
 FT /note= "fragment specifically claimed for"
 FT Region 19..738
 FT /note= "fragment specifically claimed for"
 FT Region 19..805
 FT /note= "fragment specifically claimed for"
 FT Region 133..542
 FT /note= "fragment specifically claimed for"
 FT Region 344..542
 FT /note= "fragment specifically claimed for"
 FT Region 371..402
 FT /note= "fragment specifically claimed for"
 XX
 WO200070032-A1.
 XX
 PD 23-NOV-2000.
 XX
 PD 03-MAY-2000; 2000WO-US11932.
 XX
 PR 13-MAY-1999; 99US-0311482.
 PR 27-AUG-1999; 99US-0384706.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
 DR WPI; 2001-025018/03.
 DR N-PSDB; AAC84370.
 XX
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 PT associated with inflammation such as arthritis and enterocolitis -
 XX
 PS Claim 7; Page 113-118; 125pp; English.

CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.

XX Sequence 555 AA;

Query Match 74.9%; Score 2897; DB 22; Length 555;
 Best Local Similarity 99.8%; Pred. No. 5,2e-248;
 Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEEQAKTFLDKFNEHAEEDLFYQSSLASWNTNTTEENVQNNMAGDKWSAFLEKQST 60
 DB 19 STEEQAKTFLDKFNEHAEEDLFYQSSLASWNTNTTEENVQNNMAGDKWSAFLEKQST 78
 QY 61 LAQMPFLOEIONLVKQLQALQOONGSVYSEDSKRLNTLTMTSTYTGKVCNPDNP 120
 DB 79 LAQMPFLOEIONLVKQLQALQOONGSVYSEDSKRLNTLTMTSTYTGKVCNPDNP 138
 QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKOLRELYEYVVLKNEMARAHYED 180
 DB 139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKOLRELYEYVVLKNEMARAHYED 198
 QY 181 YGDTWRGDEYVNGVDGYDSRGQLEDEYHTEFEIKPLYEHLAAVYRAKLMAVPSYISP 240
 DB 199 YGDTWRGDEYVNGVDGYDSRGQLEDEYHTEFEIKPLYEHLAAVYRAKLMAVPSYISP 258
 QY 241 IGCPLPAHLGDMWGFEWNTNLSLVTPFGQKNIDVTDAWDAQADARIFREAKEKPFVSV 300
 DB 259 IGCPLPAHLGDMWGFEWNTNLSLVTPFGQKNIDVTDAWDAQADARIFREAKEKPFVSV 318
 QY 301 GLPNMTQGFWEWSMLTDPGNVQKAVCHPTAMDLSKGFRLIMCTKYVMDDFLTAAHHEMGH 360
 DB 319 GLPNMTQGFWEWSMLTDPGNVQKAVCHPTAMDLSKGFRLIMCTKYVMDDFLTAAHHEMGH 378
 QY 361 IQYDMAVAAQPFLLRNGANGFHEAVGEIMSLSAATPKHLKSGILSPDQEDNETEINF 420
 DB 379 IQYDMAVAAQPFLLRNGANGFHEAVGEIMSLSAATPKHLKSGILSPDQEDNETEINF 438
 QY 421 LKQALTIIVGTLPTFLYMLEKRWAVFGEIPKQDMKKWEMKREIVGVVEPRPHDETVC 480
 DB 439 LKQALTIIVGTLPTFLYMLEKRWAVFGEIPKQDMKKWEMKREIVGVVEPRPHDETVC 498
 QY 481 DPASLHFVNSDYSFIRYTRTLTYOFQFOBALCOAHNEGRPLKHCIDISNSTEAGOKL 536
 DB 499 DPASLHFVNSDYSFIRYTRTLTYOFQFOBALCOAHNEGRPLKHCIDISNSTEAGOKL 554

RESULT 11

ID AAY67311 standard; Protein; 480 AA.

XX AAY67311;

XX 11-APR-2000 (first entry)

DE Human MPR015 amino acid sequence #2.

KW MPR015; treatment; hypertension; human; myocardial disease; apoplexy;

KW heart disease; apoplexy; heart disease; nervous denaturation; hormone;
 KW Alzheimer's disease; cytokine.

XX Homo sapiens.

PN JP11318472-A.

XX 24-NOV-1999.

PF 22-JAN-1999; 993P-0014949.

XX 13-MAY-1998; 98GB-0010373.

PR 18-AUG-1998; 98GB-0018009.

PA (SMIK) SMITHKLINE BEECHAM PLC.

DR WPI; 2000-109268/10.

XX MPR015 polypeptide and MPR015 polynucleotides - useful for the
 PT treatment of hypertension, myocardial diseases, apoplexy, heart
 PT diseases, nervous denaturation, Alzheimer's disease etc.

PS Claim 19; Page 20-21; 22pp; Japanese.

CC This is amino acid sequence #2 of human MPR015. The MPR015
 CC polynucleotide and polypeptide sequences can be used for the treatment of
 CC hypertension, myocardial diseases, apoplexy, heart diseases, nervous
 CC denaturation, Alzheimer's disease and diseases related to the processing
 CC of peptide hormones and cytokines.

SQ Sequence 480 AA;

Query Match 64.7%; Score 2505; DB 21; Length 480;
 Best Local Similarity 100.0%; Pred. No. 2,6e-213;
 Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEEQAKTFLDKFNEHAEEDLFYQSSLASWNTNTTEENVQNNMAGDKWSAFLEKQST 60
 DB 18 STEEQAKTFLDKFNEHAEEDLFYQSSLASWNTNTTEENVQNNMAGDKWSAFLEKQST 77
 QY 61 LAQMPFLOEIONLVKQLQALQOONGSVYSEDSKRLNTLTMTSTYTGKVCNPDNP 120
 DB 78 LAQMPFLOEIONLVKQLQALQOONGSVYSEDSKRLNTLTMTSTYTGKVCNPDNP 137
 QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKOLRELYEYVVLKNEMARAHYED 180
 DB 138 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKOLRELYEYVVLKNEMARAHYED 197
 QY 181 YGDTWRGDEYVNGVDGYDSRGQLEDEYHTEFEIKPLYEHLAAVYRAKLMAVPSYISP 240
 DB 198 YGDTWRGDEYVNGVDGYDSRGQLEDEYHTEFEIKPLYEHLAAVYRAKLMAVPSYISP 257
 QY 241 IGCPLPAHLGDMWGFEWNTNLSLVTPFGQKNIDVTDAWDAQADARIFREAKEKPFVSV 300
 DB 258 IGCPLPAHLGDMWGFEWNTNLSLVTPFGQKNIDVTDAWDAQADARIFREAKEKPFVSV 317
 QY 301 GLPNMTQGFWEWSMLTDPGNVQKAVCHPTAMDLSKGFRLIMCTKYVMDDFLTAAHHEMGH 360
 DB 318 GLPNMTQGFWEWSMLTDPGNVQKAVCHPTAMDLSKGFRLIMCTKYVMDDFLTAAHHEMGH 377
 QY 361 IQYDMAVAAQPFLLRNGANGFHEAVGEIMSLSAATPKHLKSGILSPDQEDNETEINF 420
 DB 378 IQYDMAVAAQPFLLRNGANGFHEAVGEIMSLSAATPKHLKSGILSPDQEDNETEINF 437
 QY 421 LKQALTIIVGTLPTFLYMLEKRWAVFGEIPKQDMKKWEMKREIVGVVEPRPHDETVC 480
 DB 438 LKQALTIIVGTLPTFLYMLEKRWAVFGEIPKQDMKKWEMKREIVGVVEPRPHDETVC 498

RESULT 12

ID AAU09102 standard; Protein; 261 AA.

XX

AC AAU09102;
 XX
 DT 20-DEC-2001 (first entry)
 XX
 DE Novel human protein NHP #11.
 XX
 KW Human: novel human protein; NHP; antidiabetic; antirheumatic;
 KW antiarthritic; cytosolic; antiarteriosclerotic; vulnerrary;
 KW neuroprotective; neurotropic; antiparkinsonian;
 KW anti-human immunodeficiency virus; antistimatic; vasotropic; cardiant;
 KW hypotensive; anorectic; antileptility; neuroleptic; anticonvulsant;
 KW antitumor; immunosuppressive; cerebroprotective; antileukobial;
 KW antinflammatory; antibacterial; antiproliferative; thyromimetic;
 KW immunomodulator; antiseborrheic; dermatological; vasoconstriction;
 KW gastrointestinal disorder; cardiovascular disorder; hypertension;
 KW coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;
 KW cachexia; male infertility; impotence; testicular cancer; lung tumour;
 KW hyperproliferative disorder; pulmonary system disorder;
 KW central nervous system disorder; bone disorder;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW Huntington's disease; schizophrenia; mania; dementia; paranoia;
 KW panic disorder; learning disability; amyotrophic lateral sclerosis;
 KW psychosis; autism; sleep disorder; immune system disorder;
 KW Hashimoto's thyroiditis; musculo-skeletal system disorders;
 KW diabetes mellitus; immunological disorder; asthma; AIDS; immunogen;
 KW acquired immunodeficient syndrome; leukaemia; Rheumatoid arthritis;
 KW inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;
 KW neural system disorder; respiratory disorder; olfactory disorder;
 XX wound healing.
 XX
 OS Homo sapiens.
 XX
 PN WO2001/4896-A1.
 PD 11-OCT-2001.
 XX
 PF 02-APR-2001; 2001WO-US10542.
 XX
 PR 03-APR-2000; 2000US-194118P.
 PR 29-SEP-2000; 2000US-236384P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Moore PA, Ni J, Soppet DR, Coleman TA, Gentz RL, Endress GA;
 PI Li Y, Dillon PJ;
 DR WPI: 2001-626394/72.
 DR N-PSDB: AAS14890.
 XX
 PT New human proteins, useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cardiovascular
 PT disorders, autoimmune disorders and reproductive disorders
 XX
 PS Claim 11; Page 311-312; 318pp; English.

CC brain injury and/or stroke), infectious diseases, diabetes mellitus,
 CC immunological disorders (e.g. asthma, acquired immunodeficient syndrome
 CC (AIDS), leukaemia, Rheumatoid arthritis, inflammatory bowel disease,
 CC sepsis, acne, psoriasis and lupus erythematosus), neural system
 CC disorders, respiratory disorders, olfactory disorders and wound
 CC healing. The present sequence represents an NHP of the invention.
 XX
 SQ sequence 261 AA;
 Query Match 35.1%; Score 1359; DB 22; Length 261;
 Best local similarity 99.6%; Pred. No. 4.9e-112;
 Matches 252; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 44 MNNGDWSAFLEKQSTLAQMTPLQETQNTYKQLQALQONSSVISEKSKRLTILN 103
 DB 1 MNNGDWSAFLEKQSTLAQMTPLQETQNTYKQLQALQONSSVISEKSKRLTILN 60
 OY 104 TMSITSTGKVCNPDNDOECILLEPGINEIMANSIDYNERLMMESRSVGRQLRPLYE 163
 DB 61 TMSITSTGKVCNPDNDOECILLEPGINEIMANSIDYNERLMMESRSVGRQLRPLYE 120
 OY 164 EYVLAKEMARANHEDYGDYWRGDEVNGVDYDSRGQULIEDVETFEIRPLYEHLA 223
 DB 121 EYVLAKEMARANHEDYGDYWRGDEVNGVDYDSRGQULIEDVETFEIRPLYEHLA 180
 OY 224 AYVAKLMNAPVSTISIGCLPAHLGDMMGRRTNLSITVPGQKPNIDVTDAVDOA 283
 DB 181 AYVAKLMNAPVSTISIGCLPAHLGDMMGRRTNLSITVPGQKPNIDVTDAVDOA 240
 OY 284 WDAQRIFFEAEKE 296
 DB 241 WDAQRIFFEAEKE 253
 RESULT 13
 ID AAR10426 standard; Protein: 732 AA.
 XX
 AC AAR10426;
 XX
 DT 10-APR-1991 (first entry)
 XX
 DE Human testicular angiotensin conversion enzyme.
 XX
 KW human testicular angiotensin conversion enzyme; ACE;
 KW male sterility.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= signal peptide
 FT Protein 22..732
 FT /label= mature TACE
 XX
 PN W09100354-A.
 PD 10-JAN-1991.
 XX
 PF 05-JUL-1990; 90WO-FR00513.
 PR 05-JUL-1989; 89FR-0009062.
 XX
 PA (INNM) INST NAT SANTE RECH.
 PI Soubrier F, Albenc-Gelas F, Hubert C, Corvol P;
 DR WPI: 1991-036748/05.
 DR N-PSDB: AA010328.
 XX
 PT Nucleic acid - encoding human testicular angiotensin conversion
 PT enzyme, used e.g. for in vitro detection of enzyme in organism
 XX

PS Claim 1; Fig 1; 48pp; French.

XX A bank of human testicular cDNA in Lambda gtl1 was screened with a
CC probe containing the final 3248 nucleotides of endothelial ACE. The
CC complete sequence of tACE was reconstructed from 4 separate clones.
CC The isolated nucleic acid sequence was inserted into a plasmid for
CC expression of the protein. The invention covers polypeptides
CC containing all or part of tACE sequence. These are useful in
CC treatment of inflammation or infectious diseases, especially acute
CC pancreatitis, or diseases in which kinins are involved. Antibodies
CC against the polypeptides are useful as immunoassay reagents for
CC tACE.

XX Sequence 732 AA:

Query Match 34.6%; Score 1337; DB 12; Length 732;
Best Local Similarity 41.7%; Pred. No. 2,2e-109;
Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

QY 2 TIEEQATFLDKFNEHEDLFYQSSLASMTNTNITE-----NQVNNNAGDKWSA 53
DB 70 TDEAASKFVEEYDRTQVWVNEAEANMNTNTTETSKILLQKNQJANHPT----- 123
QY 54 FLKEOSTLAOMYPLQEIQNLTVKLOLQALQNGSSVSEDKSKRLNTLNTNSTYSGK 113
DB 124 --LKGTQARKRDVNOJLQNTTITKRIKKVODLERALPAQLEENKILDMETTSVAT 181
QY 114 VCNPDNPOCELLPEGLNEIMANSIDYNERIAMESWSEVSGKOLRPLYEERYVLKNEMA 173
DB 182 VCHPNG--SCLOLEPDLTNVMAATSRKYEDLLMAMEGRDAGRAILLQFYKYVELLNQAA 239
QY 174 RANHEDYGDYKRGDYEYNGVDGIDYSRGQLEDVEHTFEIKPLYEHLAIVAKLMAA 233
DB 240 RLMGNYVAGSGWSRMYETPSLE-----ODLERLFEQLEPLYLNLHAYVRAALHHR 289
QY 234 Y-PSYISPIGCLPRAHLGDMGRFNTLYSLTVPFGQKPIIDTDAVQDAQRIEKE 292
DB 290 YGAOHILDEGPRIAHLLGNMAQTSNIYDLYVPFSPAPMDTTEMLKGMTPRMEFE 349
QY 293 AEKFEVSVGLPNTQGFENSMILDPGNVOKAVCHPTAMDLAGK-DEKILMCTKVTMDP 351
DB 350 ADDEFTSLGLLPVPEPFMNSMLEKPTDGREVVCHASAMDFYNGKDERIKQCTVMLEDL 409
QY 352 LTAHNEGHIOYMAAQAOPFLRNGANEFHNAVEGIMLSAATRKHLKSTGLSPDQ 411
DB 410 VVAHNEGHIOYMQYKDLRVALREGANPEFHAIQGVYALSVSTRHLSMLLSSEG 469
QY 412 EDNETEINFLLKQALTVGLTPTYMLEKRMVVEGGEIKPDQMKMKWEMKEIYGYE 471
DB 470 SD-EHIDNIFLMAKALDKIAIFPSYLDQKRMVFPDGSITKENTQDMSLRKYGLCP 528
QY 472 PVPHDETYCDPASLFHVSNDYSFIRYTRTLYQFOEALCQAAKHEGLAKDISNTE 531
DB 529 PVPRTOGDFFPKAKFHPSVPIRYFVSFIIOFHEALCQAAGHTGPLHKDDIYSKE 588
QY 532 AGOKLEFMAALGSEPTLLENVYAKNNVPLANTPEPLTWLKDQK--NSVVGW- 588
DB 589 AGORLATVAKMLGSRPEAMOLITGQPNMSASAMLSYKPLIDMDTEMLGEXIKGP 648
QY 589 STDMSPYADQS 599
DB 649 QYNWTPMSARS 659

RESULT 14
AAO20501
ID AAO20501 standard; Protein; 1265 AA.

XX AAO20501;
AC AAO20501;
DT 27-JUN-2002 (first entry)
XX Protein of APP related human homologue hcp51674.

XX Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;
KW amyloid precursor protein; tissue-specific expression control; human APP;
KM APP pathway modulator; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 38 /label= unknown

FT Misc-difference 39 /note= "Encoded by TGN"

FT /label= unknown

PD /note= "Encoded by CNA"

PN W0200226820-A2.

01-OCT-2001; 2001WO-BP11345.

29-SEP-2000; 2000US-236839P.

14-JUN-2001; 2001US-298309P.

(NOVS) NOVARTIS AG.

(NOVS) NOVARTIS-ERRINDONGEN VERW GES MBH.

Cohen D, Dengler UJ, Finelli AL, Freuler F, Konsolaki M;
Reinhardt MWM, Zusan S;
WPI; 2002-315796/35.

N-PSDB; AAK93935.

PT New transgenic fly, containing DNA encoding an Abeta portion of human
APP, useful for identifying agents which modulate the APP pathway and
PT which can be used to treat Alzheimer's disease -

PS Example 4; Page 94-97; 129pp; English.

XX The invention relates to a transgenic fly whose genome comprises DNA
CC encoding a polypeptide having the Abeta portion of human amyloid
CC precursor protein (APP), fused to a signal sequence. The DNA sequence
CC encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in
CC the specification. The DNA sequence is operably linked to a tissue-
CC specific expression control sequence. Expression of the sequence gives
CC the fly an altered phenotype. The purpose of the invention is for
CC identifying agents that inhibit or promote the expression and/or function
CC of genes or encoded polypeptides which modify the APP pathway. The agent
CC is a compound, triple helix DNA, antisense oligonucleotide, double
CC stranded RNA molecule, ribozyme, or particularly an antibody. It is used
CC to treat conditions such as Alzheimer's disease. The agent can be used as
CC an APP pathway modulator or in gene therapy. This sequence represents the
CC protein of the APP related human homologue hcp51674.

SO Sequence 1265 AA:

Query Match 34.6%; Score 1337; DB 23; Length 1265;
Best Local Similarity 41.7%; Pred. No. 5,2e-109;
Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

QY 2 TIEEQATFLDKFNEHEDLFYQSSLASMTNTNITE-----NQVNNNAGDKWSA 53
DB 603 TDEAASKFVEEYDRTQVWVNEAEANMNTNTTETSKILLQKNQJANHPT----- 656
QY 54 FLKEOSTLAOMYPLQEIQNLTVKLOLQALQNGSSVSEDKSKRLNTLNTNSTYSGK 113
DB 657 --LKGTQARKRDVNOJLQNTTITKRIKKVODLERALPAQLEENKILDMETTSVAT 714
QY 114 VCNPDNPOCELLPEGLNEIMANSIDYNERIAMESWSEVSGKOLRPLYEERYVLKNEMA 173
DB 715 VCHPNG--SCLOLEPDLTNVMAATSRKYEDLLMAMEGRDAGRAILLQFYKYVELLNQAA 772
QY 174 RANHEDYGDYKRGDYEYNGVDGIDYSRGQLEDVEHTFEIKPLYEHLAIVAKLMAA 233

```

Db      773 KUNGVDADSDMSRSTETPSLE-----QDERLFOELQPLYLNLHAAYRALHRH 822
Qy      234 Y-PSYISPIGCLPAHLIDGMGRFWTNLYSLTVPGOKENIDVTDAVDAQDAORIFE 292
      823 YGAGHINLEGPTRPAHLILGMMNAQWTSNIYDLVPPSPASMDTTEMLKOGTPRRMRE 882
Qy      293 AEKFFVSGLPNNTOGFENSMILTDGNNVOKAVCHPTADLKG-DFTILMCTKVTMDP 351
      883 ADDEFTSLGLLPVPEPFNKSMLKEPTDGEVYVNASAMDFNGKDFRIKCTTVNLDEL 942
Qy      352 LTAHHEHGHIOYDMAAOPFLIRNGANGFHEAVGEIMSLASAPKHLKSLGSLSPDRQ 411
      943 VVAHHEHGHIOYFMQYKDLVVALREGANPGFHEALIGVLAISVTPKHLHSLNLSSEGG 1002
Qy      412 ENNETEINFLKQALITVGLPFTYMLEKRMWVFKGEIPKQDMKKWMEKREIYVVE 471
      1003 SD-EHNDINFLMKALDKIAIFPSYLVQDMRWVRVDSITKENYNOEMWSLAKYQGLCP 1061
Qy      472 PVPHDETCDPASLHVNDYSFIRYITTYLYQFOFQALCOAHHEGPLKCDISNTE 531
      1062 PVPRTOGDDPFAKFIHPSVPIRYFVSFIIOFQFHEALCOAGHGTGLHKCDIYQSK 1121
Qy      532 ACQKLENNLRIGKSEPTLALENVYGAKNMNYRPLNPEPLFTWIKQONK--NSFGW- 588
      1122 AGORLATAMKLGFSRWRPAMQITQPMASAMLSIFKPLLDMLKTENELHGEKLGMP 1181
Qy      589 STDMSPYADQS 599
      1182 QYNTWTPNSARS 1192
Db

```

RESULT 15
AA04111
ID AA04111 standard; peptide; 1306 AA.

AC AA04111;
XX
DT 07-SEP-1990 (first entry)
XX
DE Human angiotensin converting enzyme (ACE).
XX
KW human angiotensin converting enzyme; hypertension; bradykinin.
XX
OS synthetic.

Key
FH Location/Qualifiers
FT 30..1277
FT /label-mature ACE
FT /note="derived from pre-ACE by removal of signal peptide"
FT Modified-site
FT 38..38
FT /label-putative N-glycosylation site
FT 54..56
FT /label-putative N-glycosylation site
FT 74..76
FT /label-putative N-glycosylation site
FT 111..113
FT /label-putative N-glycosylation site
FT 146..148
FT /label-putative N-glycosylation site
FT 160..162
FT /label-putative N-glycosylation site
FT 318..320
FT /label-putative N-glycosylation site
FT 445..447
FT /label-putative N-glycosylation site
FT 509..511
FT /label-putative N-glycosylation site
FT 523..525
FT /label-putative N-glycosylation site
FT 677..679
FT /label-putative N-glycosylation site
FT 713..715
FT Modified-site

```

FT      760..762 /label-putative N-glycosylation site
FT      942..944 /label-putative N-glycosylation site
FT      1191..1193 /label-putative N-glycosylation site
FT      1225..1227 /label-putative N-glycosylation site
FT      1225..1227 /label-putative N-glycosylation site
Qy      Modified-site
      1225..1227 /label-putative N-glycosylation site
Db      WO9003435-A.
      05-APR-1990.
      27-SEP-1989; 89WO-FR00496.
      27-SEP-1988; 88FR-0012620.
      (INRM) INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE.
      Soubrier F, Alhenc-Gelas F, Hubert C, Corvol P;
      WPI: 1990-132272/17.
      N-PSDB: AA004027.
      DR
      CC Human DNA encoding human angiotensin converting enzyme used eg in
      CC diagnosis of hypertension, evaluation of enzyme inhibitors
      CC Disclosure: ; p; French.
      PS
      XX Human angiotensin converting enzyme hydrolyses angiotensin I and kinins.
      CC Either intact enzyme or fragments thereof can be used to generate
      CC antibodies for diagnostic use. Oligonucleotide probes can also be made
      CC which are complementary to the sequence encoding the enzyme.
      SQ
      Query Match 34.6%; Score 1337; DB 11; Length 1306;
      Best Local Similarity 41.7%; Pred. No. 5.5e-109;
      Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

```

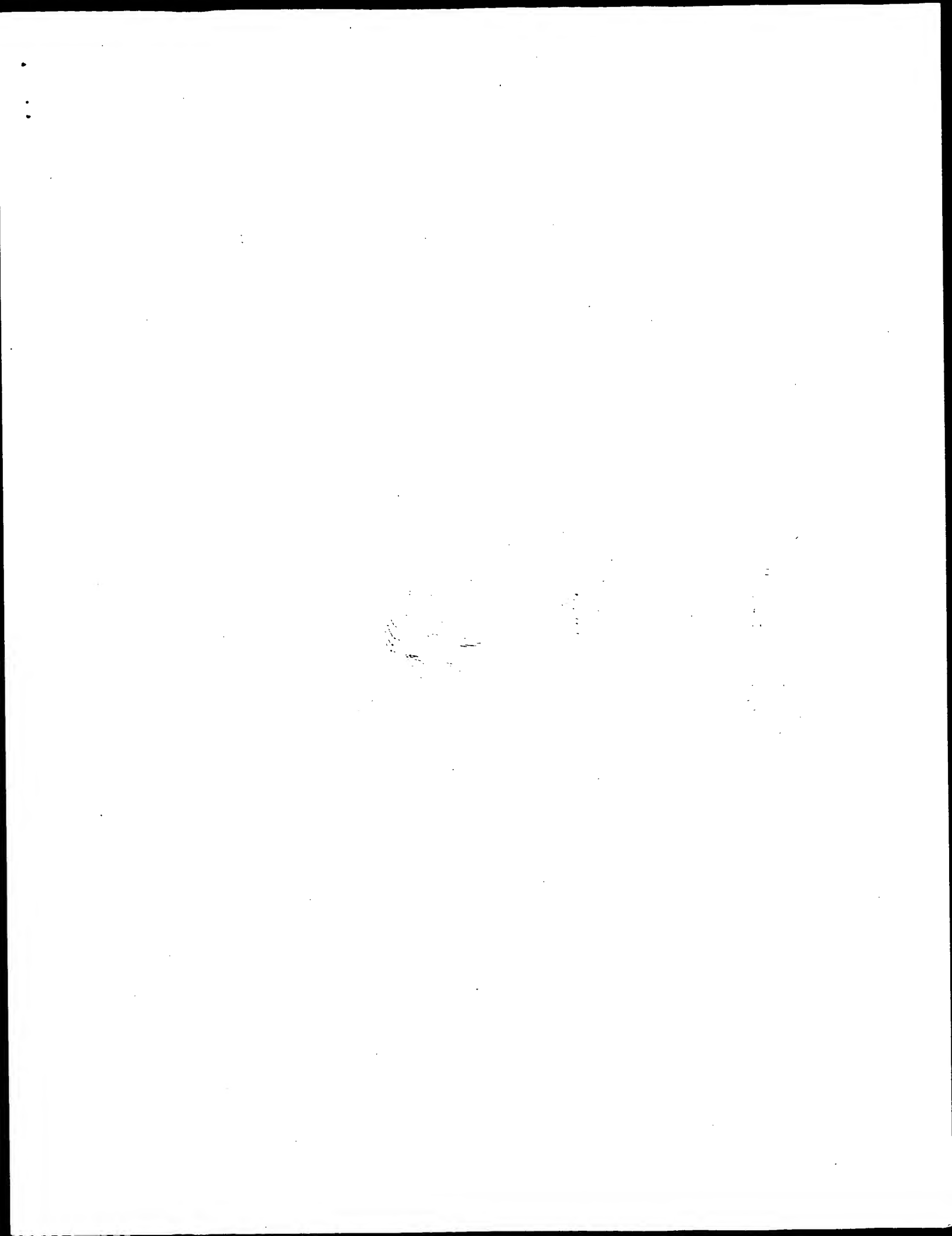
2 TIEQAKTFLDKFHEADELFYQSSLASWYNTNITE-----NVQNMNNAQDKWSA 53
DB IDPEAKSFVEHYDRTQVWVNEAYEAANWYNTNITETSKILLQKMOJANTHT----- 697
Qy 54 FLKEOSTLAQMYPLQEIQNLTVKLOALQONSSVSEKSKRLNTILTMSTIYSTGK 113
DB 698 --LKGTOARKFTVNOQNTTIKRIKYODLEALPAQDELEYKTLIDMETTYSVAT 755
Qy 114 VCPNDPQECILIEPGINEIMANSLDYNERIMAWESRSEVQGLRPLYEYVVLNENMA 173
DB 756 VCHPNG--SCIOLEPDLTNVATSRKEYEDLLMAEGRDKAGRAILLQFYKYELLNOQA 813
Qy 174 RANVEYEGDWRDGVNVDGVDSRGQLEDEVTPEPEIRYVYENHLAAVYRAKLMA 233
DB 814 KUNGVDADSDMSRSTETPSLE-----QDERLFOELQPLYLNLHAAYRALHRH 863
Qy 234 Y-PSYISPIGCLPAHLIDGMGRFWTNLYSLTVPGOKPRIDVTDAVDAQDAORIFE 292
DB 864 YGAGHINLEGPTRPAHLILGMMNAQWTSNIYDLVPPSPASMDTTEMLKOGTPRRMRE 923
Qy 293 AEKFFVSGLPNNTOGFENSMILTDGNNVOKAVCHPTADLKG-DFTILMCTKVTMDP 351
DB 924 ADDEFTSLGLLPVPEPFNKSMLKEPTDGEVYVNASAMDFNGKDFRIKCTTVNLDEL 983
Qy 352 LTAHHEHGHIOYDMAAOPFLIRNGANGFHEAVGEIMSLASAPKHLKSLGSLSPDRQ 411
DB 984 VVAHHEHGHIOYFMQYKDLVVALREGANPGFHEALIGVLAISVTPKHLHSLNLSSEGG 1043
Qy 412 ENNETEINFLKQALITVGLPFTYMLEKRMWVFKGEIPKQDMKKWMEKREIYVVE 471
DB 1044 SD-EHNDINFLMKALDKIAIFPSYLVQDMRWVRVDSITKENYNOEMWSLAKYQGLCP 1102

KM vascular disorder; asbestosis.
 XX
 OS Mus sp..
 XX
 PN WO200136632-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000WO-IL00766.
 XX
 PR 17-NOV-1999; 99IL-0132978.
 PR 10-DEC-1999; 99IL-0133455.
 XX
 PA (COMP-) COMPUGEN LTD.
 XX
 PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 DR WPI: 2001-336004/35.
 DR N-PSDB: AAS06057.
 XX
 PT Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies
 XX
 PS Claim 4; Fig 57; 519pp; English.
 XX
 CC The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen p53, and vasocactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, noncolidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis.
 CC
 CC Sequence 1249 AA;
 SQ
 Query Match 34.5%; Score 1334; DB 22; Length 1249;
 Best Local Similarity 42.6%; Pred. No. 9.5e-109;
 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;
 QY 2 TIEQAKTFLDKNEHAEDELVYQSSIASWNTNTITEENVONNANAGDKMSAFLEKOSTL 61
 DB 649 TDEAKADREVEEDRTAOVLINLEYAANMOYNTNITEGSKILLERSTEVSHHTLYGTR 708
 QY 62 AOMYPILOEIONLTVKIQLOALQONGSSVLEDSKRLNTLNTMTSTYSGKVCNPDNQ 121
 DB 709 AKTFDVSNFONSSIKRIKILQNLDRAYLPKLELEYNOILLDMETTYLSNICYNG-- 766
 QY 122 ECLLEFGLNEIANSIDLYNERLWAMESRSEVKOLRPLYEEVYVILKNEMARANYEDY 181
 DB 767 TCMPLPEDLTNMATSKEELLMAMKSRDKYGRALLPFPKRYVESNRIATLNGTDA 826
 QY 182 GDVWRGDYEVNGVDGYISRGOLIEVDEHFEELKPLYEHLAVYRAKLMNAPVSYSP 240
 DB 827 GDSNRSLEYEDNLE-----QDLKELYOELOPLYLMLHAYVRSLSLRHNGSEYINL 876
 QY 241 IGCPLAHLIDMNGREFNTNLSLTPFGQKPNIDVDVAVQADADRIKFEKAEFEVSV 300
 DB 877 DGPFLAHLIDMNGREFNTNLSLTPFGQKPNIDVDVAVQADADRIKFEKAEFEVSV 300
 QY 301 GLPMTGCEWENSMALDPCGVAVCAVCHPTAMDJGK--DFRILMCTKTKTMDFLAHHEMG 359
 DB 937 GLPMTGCEWENSMALDPCGVAVCAVCHPTAMDJGK--DFRILMCTKTKTMDFLAHHEMG 359

QY 360 HIQYDMAVAAQPELLNGANEGFHEAVGEIMLSAATPKHLKSLIGLLSPDQENDETEIN 419
 DB 997 HIQYFMQYNDLPVTEFEGANPGFHEAIDIMALSTPKHLKSLINLSTPE--GSGYEYDIN 1055
 QY 420 FLKQALITVGLTFPFTYMEKRWKWFKEGELPKRDKMKKWKMKREIVGVPEPHDETY 479
 DB 1056 FLKQALITVGLTFPFTYMEKRWKWFKEGELPKRDKMKKWKMKREIVGVPEPHDETY 479
 QY 480 CDPASLFHVSNDYSFRTYRTLYQFQFOALCOAQBHGPPLKDKDISNTEAGOKLFEM 539
 DB 1116 FDPGSKFHPANPYRYRFFSFIQFQFHALCRAAGHTGPKHCKDIYQSKRAGKLLADA 1175
 QY 540 LRLKSEPTLALENVGAKNANVRPLNFEPLFTWLDQNK--NSFYGM--STWSP 594
 DB 1176 MKLGSKFHPANPYRYRFFSFIQFQFHALCRAAGHTGPKHCKDIYQSKRAGKLLADA 1175
 RESULT 18
 AA002985
 ID AA002985 standard; Protein: 1252 AA.
 AC AA002985;
 XX
 DE 12-SEP-2001 (first entry)
 XX
 DE Angiotensin converting enzyme (ACEV) splice variant protein #85.
 XX
 KM Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KM granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KM platelet-derived endothelial cell growth factor; cardiovascular disease;
 KM cellular tumour antigen p53; cyclin-dependent kinase inhibitor 1C;
 KM vasocactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KM myocardial infarction; coronary arterial thrombosis; renal disease;
 KM diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KM multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KM noncolidotic pulmonary granulomatous disease; endothelial abnormality;
 KM vascular disorder; asbestosis.
 KM
 OS Mus sp..
 XX
 PN WO200136632-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000WO-IL00766.
 XX
 PR 17-NOV-1999; 99IL-0132978.
 PR 10-DEC-1999; 99IL-0133455.
 XX
 PA (COMP-) COMPUGEN LTD.
 XX
 PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 DR WPI: 2001-336004/35.
 DR N-PSDB: AAS06085.
 XX
 PT Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies
 XX
 PS Claim 4; Fig 85; 519pp; English.
 XX
 CC The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen p53, and vasocactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various

Oy		535 KLENNRLRCKSEPTTALENVYGAKNMNVRPLANFEPLFTWKLQN--KNSFGWST 550
Dd		: : : : : : : : : : 551 AFHNLMSGASRPWDALFAFNGERIMSGKATAEYFEPLRWLAEENIKNNVIGWT 608
RESULT 20		
AAR70013		
ID	AAR70013 standard; Protein; 660 AA.	
XX AC AAR70013;		
DT	25-SEP-1995 (first entry)	
XX DE	Tick carboxypeptidase.	
XX KM	Tick; antigen; carboxypeptidase; vaccine.	
OS	Boophilus microplus.	
XX FH	Key	Location/Qualifiers
FT	Peptide	30..53
FT	Peptide	/label= BM91 peptide
FT	Peptide	185..201
FT	Peptide	/label= peptide AAT91251
FT	Peptide	202..209
FT	Peptide	/label= AAT91141
FT	Peptide	225..253
FT	Peptide	/label= T9118
FT	Peptide	262..269
FT	Peptide	/label= Bm91 peptide
FT	Peptide	442..452
FT	Peptide	/label= T9129(a)
FT	Peptide	456..481
FT	Peptide	/label= T9129(b)
FT	Peptide	599..606
FT	Peptide	/label= T9109
FT	Domain	1..29
FT		/label= signal
FT		639..655
FT		/label= C-terminal transmembrane domain
FT		/note= "putative"
PD	WO9504827-A.	
XX PN	16-FEB-1995.	
XX PD		
XX PF	10-AUG-1994; 94WO-AU00463.	
XX PR	10-AUG-1993; 93AU-0000458.	
XX PA	(BIOT-) BIOTECH AUSTRALIA PTY LTD.	
PA	(CSIR) COMMONWEALTH SCI & IND RES ORG.	
Pt	Cobon GS, Kemp DH, Teliam RL, Willadsen P;	
XX DR	WPJ; 1995-090905/12.	
DR N-PSDB; AAQ82948.		
XX	New antigenic tick carboxypeptidase and corresp. DNA - are used	
Pt	in vaccines for producing antibodies against ticks, insects and	
Pt	nematodes.	
XX	Disclosure; Fig 6; 138pp; English.	
XX	Clone A5 was prep'd. from adult tick cDNA library. Clone 4U1 was	
CC	prep'd. from the larval stage of B. microplus (Calliope strain).	
CC	AAQ82948 is a hybrid of sequences from clone 4U1 (nt 1-966 & 1747-2047)	
CC	and A5 (nts 967-1746). The translation of the tick carboxypeptidase	
CC	cDNA sequence is shown in AAR70013. All the native tick	
CC	carboxypeptidase sequences listed in Table 11 (see AAR70014-R70023)	
CC	are found in the translation (see FT). The predicted aa sequence	
CC	agrees with the peptide sequence for all peptides w/tn 2 exceptions.	
CC	These differences are Asp for Gln14 in peptide T9126, and Asn for	

[illegible]



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

May 26, 2003, 17:53:54 ; Search time 22 Seconds
(without alignments)
962.932 Million cell updates/sec.

Title: US-09-978-385-2_COPY_19_738
Perfect score: 3869
Sequence: 1 STIEQAKTFIDKRNHAEI.....DNSLEFLGIQPTLGPNOFP 720

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	805	4	US-08-989-299-2
2	1337	34.6	732	1	US-08-481-626-2
3	1337	34.6	732	4	US-08-989-299-4
4	1337	34.6	1306	4	US-08-989-299-7
5	1334	34.5	732	4	US-08-989-299-5
6	1334	34.5	1312	4	US-08-989-299-8
7	1310	33.9	1313	4	US-08-989-299-6
8	1283	33.2	737	4	US-08-989-299-10
9	1283	33.2	1310	4	US-08-989-299-11
10	1073	27.7	615	4	US-09-440-325A-1
11	989	25.6	907	4	US-08-989-299-12
12	638.5	16.5	907	4	US-08-989-299-15
13	120	3.1	990	2	US-08-630-916A-46
14	116	3.0	990	2	US-08-392-625-20
15	115	3.0	990	2	US-08-446-059-2
16	113.5	2.9	665	3	US-08-844-059-2
17	113.5	2.9	665	4	US-09-431-202-2
18	113.5	2.9	665	4	US-08-491-357-2
19	104	2.7	834	1	US-08-968-633-2
20	104	2.7	834	3	US-09-196-466-2
21	104	2.7	834	5	PCT-US96-10823-2
22	104	2.7	631	4	US-08-448-489-17
23	102.5	2.6	660	3	US-08-704-711A-18
24	102.5	2.6	660	4	US-09-521-220-18
25	102.5	2.6	660	4	US-09-391-104-19
26	102.5	2.6	660	4	US-09-453-702B-257
27	102	2.6	3169	4	Sequence 257, App

28	101.5	2.6	450	4	US-09-306-593-13	Sequence 13, Appl
29	101.5	2.6	560	4	US-09-134-001C-3153	Sequence 3153, Ap
30	99.5	2.6	930	4	US-08-953-040-2	Sequence 2, Appl
31	99	2.6	789	4	US-09-002-285-84	Sequence 84, Appl
32	97.5	2.5	570	3	US-08-747-221B-55	Sequence 55, Appl
33	97.5	2.5	570	4	US-09-005-051-55	Sequence 55, Appl
34	97.5	2.5	595	3	US-08-747-221B-31	Sequence 31, Appl
35	97.5	2.5	595	4	US-09-005-051-31	Sequence 31, Appl
36	97	2.5	977	3	US-08-335-844A-22	Sequence 22, Appl
37	96	2.5	789	4	US-09-002-285-82	Sequence 82, Appl
38	96	2.5	806	4	US-08-945-983-2	Sequence 2, Appl
39	96	2.5	3443	2	US-08-416-603-2	Sequence 2, Appl
40	94	2.4	452	1	US-08-317-880-2	Sequence 2, Appl
41	94	2.4	452	2	US-08-782-396-2	Sequence 2, Appl
42	93	2.4	647	2	US-08-844-056-2	Sequence 2, Appl
43	93	2.4	789	4	US-09-002-285-78	Sequence 78, Appl
44	93	2.4	789	4	US-09-002-285-94	Sequence 94, Appl
45	93	2.4	790	4	US-08-960-780-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-989-299-2
Sequence 2, Application US/08989299
Patent No. 6194556

GENERAL INFORMATION:

APPLICANT: Acton, Susan L.
APPLICANT: Robinson, Keith E.
TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,299
FILING DATE: 11-DEC-1997
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Arnold E., Beth
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MIA-025.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 805 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein

Query Match 100.0%; Score 3869; DB 4; Length 805;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 1 STIEQAKTFIDKRNHAEIIFYOSSIASWNTNTIENYQMMNAGDKWSAFLEKOST 60
DB 19 STIEQAKTFIDKRNHAEIIFYOSSIASWNTNTIENYQMMNAGDKWSAFLEKOST 78
0Y 61 LAQMTPLQEIQNTLVKIQIALQAGNSSVLSSEKSKRLNTILNTMTSTIYSTRKVCNPDP 120

```

Db          79  LAQWYPIQEIONTLVTKIQLQALDQNGSSVLSSEKSKRLNTLITMSTIYSTGVCVCPDNP 138
              |||
QY          121  QECILLEGLEIMANSLDIYNERLMAESWRSREYQKQRLDYEEYVLAKEMARAHYED 180
              |||
Db          139  QECILLEGLEIMANSLDIYNERLMAESWRSREYQKQRLDYEEYVLAKEMARAHYED 198
              |||
QY          181  YGDYWRGDIYVNGVDGDIYDSRGQILEVEHTEFEERIKPLIYHILAAVYAKIMNYPYSISP 240
              |||
QY          199  YGDYWRGDIYVNGVDGDIYDSRGQILEVEHTEFEERIKPLIYHILAAVYAKIMNYPYSISP 258
              |||
Db          241  IGCPLAHLLGDMGGEFTNLXSLVYPPGQKPNIDVTAMVDAQADQRIEKEAKEFVSV 300
              |||
Db          259  IGCPLAHLLGDMGGEFTNLXSLVYPPGQKPNIDVTAMVDAQADQRIEKEAKEFVSV 318
              |||
QY          301  GLPNNTOGFWENSMULTDPCNVQKAVCPHTANDLKGDFRILMCTKVTMMDFLAHHBMGH 360
              |||
QY          319  GLPNNTOGFWENSMULTDPCNVQKAVCPHTANDLKGDFRILMCTKVTMMDFLAHHBMGH 378
              |||
Db          361  IOYDMAVAQAOPRLNKGANEGRHEVAVGEIMSLAATPRKHLKSIGLSPDEQEDETEINF 420
              |||
Db          379  IOYDMAVAQAOPRLNKGANEGRHEVAVGEIMSLAATPRKHLKSIGLSPDEQEDETEINF 438
              |||
QY          421  LTKQALTYGLTPEFTYMLEKRWAVFKGEIPKQDMKKWMMKREIYGVVBPVPHDFTYC 480
              |||
Db          439  LTKQALTYGLTPEFTYMLEKRWAVFKGEIPKQDMKKWMMKREIYGVVBPVPHDFTYC 498
              |||
QY          481  DPAELFHSNDVSFIRYVTRTLVYQFOFQALCQAKHGGPLHKDISNSTEAGOKLFNML 540
              |||
Db          499  DPAELFHSNDVSFIRYVTRTLVYQFOFQALCQAKHGGPLHKDISNSTEAGOKLFNML 558
              |||
QY          541  RLGKSEBWTALLENVGAKNMVRPLNLTPEELFTWLKDQKNSVVGSTMSPTADDSI 600
              |||
Db          559  RLGKSEBWTALLENVGAKNMVRPLNLTPEELFTWLKDQKNSVVGSTMSPTADDSI 618
              |||
QY          601  KVAIISLKSALGDCAVEMNNENMYLFESSVAVAMROFTLKVKKNQMLTFGEEDRYANLKPR 660
              |||
Db          619  KVAIISLKSALGDCAVEMNNENMYLFESSVAVAMROFTLKVKKNQMLTFGEEDRYANLKPR 678
              |||
QY          661  ISFNEFTVAKNVDIIPTEVEKAIKMSRSRINDAFRLNDSLEFLGIQPLGPDPNP 720
              |||
Db          679  ISFNEFTVAKNVDIIPTEVEKAIKMSRSRINDAFRLNDSLEFLGIQPLGPDPNP 738
              |||

RESULT 2
; US-08-481-626-2
; Sequence 2, Application US/08481626
; Patent No. 5801040
; GENERAL INFORMATION:
; APPLICANT: Soubrier, Florent
; APPLICANT: Albenc-Gelas, Francois
; APPLICANT: Hubert, Chrichtine
; APPLICANT: Corvol, Pierre
; TITLE OF INVENTION: Nucleic Acid Coding for the Human
; TITLE OF INVENTION: Testicular Angiotensin Converting Enzyme (ACE) and its
; TITLE OF INVENTION: Uses, Especially for the In Vitro Screening for this
; TITLE OF INVENTION: Enzyme in the Organism
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flunegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,626
; FILING DATE: 07-JUN-1995

```

	CLASSIFICATION: 435	
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER: US 07/656,183	
	FILING DATE: 04-MAR-1991	
	CLASSIFICATION: 435	
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER: FR 89-09062	
	FILING DATE: 05-JUL-1989	
	ATTORNEY/AGENT INFORMATION:	
	NAME: Meyers, Kenneth J.	
	REGISTRATION NUMBER: 25,146	
	REFERENCE/DOCKET NUMBER: 04,958-0006-02000	
	TELECOMMUNICATION INFORMATION:	
	TELEPHONE: 202-408-4400	
	TELEFAX: 202-408-4400	
	INFORMATION FOR SEQ ID NO: 2:	
	SEQUENCE CHARACTERISTICS:	
	LENGTH: 732 amino acids	
	TYPE: amino acid	
	STRANDEDNESS: single	
	TOPOLOGY: linear	
	MOLECULE TYPE: protein	
	US-08-481-626-2	
Query Match	34.6%; Score 1337; DB 1; Length 732;	
Best Local Similarity	41.7%; Pred. No. 1,16-119;	
Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;		
QY	2 TIEQAKTFLDKHEAEDELFTYQSSLASNNYNTITEE-----NYQNMNAGDKMSA 53	
DB	70 TDEAEASKFEFEEDRISQVWNEVEAMNNNTITETSKILLQKMQLANT----- 123	
QY	54 FIKQOSTLAQMYLQEIOMLTKLOALQOQSSVLSSEDKSKRLNTILNTMSTIYSGK 113	
DB	124 --LKYGQAKKFEVNOLOMTTIKRIKKYQDLERAAIPQGLEBYNKILDMETYSVAT 181	
QY	114 VCNPDNFOECLLEPGINETMANSLDYNERIMAMESRSEVGRKRLRLYEYVLKINEMA 173	
DB	182 VCHNG--SCLQTEPDLITVMATSRKYEDLMAWGMGMDKGRALIDFYKRYVELNQAA 239	
QY	174 RANHYEDGYGWMGDVEYVNGVDGYDSRGQLIEDVEHFEFEIKPLYEHLHAYRAKLMA 233	
DB	240 KLVNGVDAQSWMSMYTETPSE-----QDLERLFOELQPLYLNLHAYRRALHRI 289	
QY	234 Y-PSYISPIGLPAHLIGDMGRFTNLVSLVPPGQKPNIDVDANVDQAQARIFE 292	
DB	290 YGAQINIEGPIPAHLIGNNMAQTWSNIYDLVPPSPASMDTYEAMLKQGWTPRMEKE 349	
QY	293 AEKRFVSGLPNMQGFWENSMITDGNVOKAVCPHFTAMDELGK--DERILMCTKVTMDF 351	
DB	350 ADDEFTSLGLPVPPEFWKSMLEKFTPDGREVYCHASAMDPEYNGKDEFIKQCTVLEDL 409	
QY	352 ITAHEHMGHIDYMAVAAPFLLRNGANBSGFHEAVGEIMLSAATPRKILKIGLSPDQ 411	
DB	410 VVAHHEMGHIOFYMQYKDLPLVALREGANPGEHIALDVALSVSPKHLHSLNLSSEGG 469	
QY	412 ENNEFEINFLKQALITVGLTPFYMLKEKRMVVEFGELIPKQDMKKMKMEKKREIVGVE 471	
DB	470 SD-EHDIWFLMKMAIDKIAITPSYILDQQRKMYVEFGSGITKKNYQEWMSLRKYQGLCP 528	
QY	472 VPVHDEYTCDPASLEFHVSNDSYSTRYTRTLVYFOFQEARLCOAANBEPILAKCOISNTE 533	
DB	529 PVPRPGQGFDEPAKKEIPSSVPIRYRVSFIIQOFHEALCOAAGHTPRLKRCIDYOSKE 586	
QY	532 AGQKLFMLKRLKGSPTWLLALENVYGVKKNVNPRLPEPTFLMKLQKNC--NSFYGV 588	
DB	589 AQRRLATAKLGFSPPWEAPMOLITIGQNSASAMLSYKRPDLMLRLRENELHGRKLGWP 646	
QY	589 STDMSPYADQS 599	
DB	649 QYWMTPNSARS 659	

```

RESULT 3
US-08-989-299-4
; Sequence 4, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Robinson, Keith E.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold E., Beth
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MIA-025.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-7000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-989-299-4

Query Match
Best Local Similarity 34.6%; Score 1337; DB 4; Length 732;
Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

QY 2 TIEBQAKTFLDKFNEHAEDELFEYSSLASMYNTNITE-----NVQNNMAGDKWSA 53
DB 70 TDEAASKVEYEYRTSOVWMEYAEANWNTNTTETSKILLQKNMIAHNT-----123
QY 54 FLKEQSTLAQMYPLQEIQNTLVKIQLOALQONGSSVLSSEDKSKRLNTIINTMSTIYSTGK 113
DB 124 --LKYGTQARKFDVNOJQNTTIKRIKKVODLERALPAGLEERYNKILIDMETTVSVAT 161
QY 114 VCPNDPQECILLEGLNEIMANSIDYNERLMAWESRSEYKQRLPYEYVVLKNEMA 173
DB 162 VCHENG--SCQLEPDLNVAATSRKYEDLIMAWEGWRDAGRAIILQFYKVELINQAA 239
QY 174 RANNEYDYGDYWRGDEYVNGVGYDSRGLIEVEHTFEIRKPLRYHLAAVRAKIMNA 233
DB 240 RLNGYVDAGDSMSRYEPPSLE-----QLELFEQLDQPLVNLHAYVRALHRR 289
QY 234 Y-PSYISPIGLPRAHLIGDMGRFTWNLKSLTVPGQKPNIDYVDAMVDQADQRIKFE 292
DB 290 YGAQININIEGPRIAPHLIGNMWAOQWNSNIYDLVVPFSPASMDTTEALVKGQWPRRME 349
QY 293 AKKFEVSVGLPNNMTOGFENSMILTDPGNVOKAVCHPAMDIGS-DFRIIMCTKYTMDF 351
DB 350 ADDEFTSLGLPVPPEPMKSMLEKPTDQREYVCHASAMDFYNGKDFRIKQCTTVLIDL 409
QY 352 LTAHHEGHIOYDMAVYAAQPELLRNGANEGFHEAVEIMSLSATPKHKSIGLISPDQ 411
DB 410 VVAHHEGHIOYFQYKDLPLVALREGANPFGFHEAIGVLAISVTKHLHSINTLSSEG 469

```

```

RESULT 4
US-08-989-299-7
; Sequence 7, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Robinson, Keith E.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold E., Beth
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MIA-025.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-7000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1306 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-989-299-7

Query Match
Best Local Similarity 34.6%; Score 1337; DB 4; Length 1306;
Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

QY 2 TIEBQAKTFLDKFNEHAEDELFEYSSLASMYNTNITE-----NVQNNMAGDKWSA 53
DB 644 TDEAASKVEYEYRTSOVWMEYAEANWNTNTTETSKILLQKNMIAHNT-----697
QY 54 FLKEQSTLAQMYPLQEIQNTLVKIQLOALQONGSSVLSSEDKSKRLNTIINTMSTIYSTGK 113
DB 698 --LKYGTQARKFDVNOJQNTTIKRIKKVODLERALPAGLEERYNKILIDMETTVSVAT 755
QY 114 VCPNDPQECILLEGLNEIMANSIDYNERLMAWESRSEYKQRLPYEYVVLKNEMA 173

```

Db	756	VCHPBG--SCIOLEBDLTVMATSRKYEIDLMMMEGRKAKGRALIQTYPRYVELINGAA	813
Qy	174	RANHEVEDGYWRGDEYENGVGDYDSRGQLIEDVEHTEEIKPLYEHLHAYRAKLMA	233
Db	814	RUNGVDADGDSRMSYETPSE-----ODLERLQELQPLYLMDHYRAALHRH	863
Qy	234	Y-PSYISPIGLIPAHLLDMMGRWTVLTLVPRFOKNDIVDTAMVDQADARIPKE	292
Db	864	YGAOHINTEGPIPAHLGNMAQAQWSNITDLYVPEFSASMDTEAMLKQGTPEAKKE	923
Qy	293	AKKEFYSVGLPMNTGFEFENSMLDPPNVOKAVCHPTANDLGK-DEFILMCTKYMDF	351
Db	924	ADDEFISLLELPVPEPFNNKMLKPRIDGREGVCHASAMDFYNGKDFRIKQCTVNLDDL	983
Qy	352	LTAHEHMGHIQYDMAYYAAQPLLRNGANEFEHVAEGEIMSAAATPKHLKSGILSPDQ	411
Db	984	VVAHHMGHIQYFMQYKDLPVALRGANPGHGAIGVYLAIVSTPKHLASINLSSGQ	1043
Qy	412	EDNETINFLKQALITVGLTPEYTMLEKRWMMYFKELIPKDDMAKRWEMREITGVYE	471
Db	1044	SD-EHINFLMKALDKLFIPEFSTVDQMRKRFDSITIKENYNGEMWSILKRYQLCP	1102
Qy	472	PVPHEDEYCDPASLFEHVSNDYSFIRYTRILYQFOFQDALCOAKHREGLHCDISNTE	531
Db	1103	PVPFGODEPFGCAKFIIPSSVPYIRYFVSFLIOQFHEALCOAGHRTGLPKCDITYOSK	1162
Qy	532	AGQKLEFMMRLIGSEFWTLATENYVGANMVRRLVYFPLETWLKDQNK--NSEFVG-	588
Db	1163	AGQALATAMALIGSRPRAAMQILITQFNRMSASMLTYFPRLLDMLTENELGERTGMP	1222
Qy	589	STDMSPYADDS 599	
Db	1223	QYNTPSARS 1233	
Db	1223	QYNTPSARS 1233	

```

1 STRANDEDNESS:
2 TOPOLOGY: 1linear
3 MOLECULE TYPE: protein
4 US-08-989-299-5
5
6 Query Match 34.5%; Score 1334; DB 4; Length 732;
7 Best Local Similarity 42.6%; Pred. No. 2.2e-119;
8 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7
9
10 QY 2 TIEQAKTFLDFENHAEDELVQSSLASMTNNTNTEENYQNNMNGDGKSAFLKEOSTL 61
11 DB TDAKADKRVVEEDRDPAQVLNEYAEANWQNTNITTEGSKILLESTEVSNLTLYGTR 128
12 QY 62 AQMPYPLQEIQNTLVKLQALQONSSVLSBDSKRLNTLTNTSTYSTGKACPNQNDQ 121
13 DB TDKAKADKRVVEEDRDPAQVLNEYAEANWQNTNITTEGSKILLESTEVSNLTLYGTR 128
14 QY 129 AKTFEDPNSNQNSIKRIKKLQNLDRALVPRELEEVYQQLIDMETYSLSNICYTG-- 186
15 DB AKTFEDPNSNQNSIKRIKKLQNLDRALVPRELEEVYQQLIDMETYSLSNICYTG-- 186
16 QY 122 ECLLEPGLNETMANSLDYNERLMAWESWRSGKOLRLPYEEVYVLKNNMABANHYEDY 181
17 DB TCMPLBEDLNNMATAKRYKEELLMAMKSRDVGKALILPEPKYVESSNKIAKNGYTDA 246
18 QY 182 GDVVRGDEYVNGDYGXDYSRGOLIEDVEHTEFEELPLEYHLAHAVRAKLMAAPYS-YISP 240
19 DB GDVVRGDEYVNGDYGXDYSRGOLIEDVEHTEFEELPLEYHLAHAVRAKLMAAPYS-YISP 240
20 QY 247 GDSKRSIYESDNLE-----QDLKLYQELQPLVTLNHAVYRSLSRHGYSEYINL 296
21 DB GDSKRSIYESDNLE-----QDLKLYQELQPLVTLNHAVYRSLSRHGYSEYINL 296
22 QY 241 IGCPLPAHLIGDMWGRFTWNLISLTPFGCKPNIQDTADMANVQADQARIFREKKEFPVS 300
23 DB IGCPLPAHLIGDMWGRFTWNLISLTPFGCKPNIQDTADMANVQADQARIFREKKEFPVS 300
24 QY 297 DGPRIAPHLILNNMAQWTMSNIYDLYAFPSAPNIDATEMIKQGTPTKRIEEDNFTSL 356
25 DB DGPRIAPHLILNNMAQWTMSNIYDLYAFPSAPNIDATEMIKQGTPTKRIEEDNFTSL 356
26 QY 301 GLPMPWQGPWNSMLNDPGVQKAVCHPAPMDLGKG-DPRILMCKYVMDDFLTAHHMG 359
27 DB GLPMPWQGPWNSMLNDPGVQKAVCHPAPMDLGKG-DPRILMCKYVMDDFLTAHHMG 359
28 QY 357 GLPVPPEPFWNKSLEKFTDGRVYVCHPSAMPDFYNGKQPRIKCTSVAMEDLYAHHEMG 416
29 DB GLPVPPEPFWNKSLEKFTDGRVYVCHPSAMPDFYNGKQPRIKCTSVAMEDLYAHHEMG 416
30 QY 360 HIOYDMAYAAQPELLRNGANEGHEAVGEINLSAATPRHLKSGTLLSPDQEDNEIN 419
31 DB HIOYDMAYAAQPELLRNGANEGHEAVGEINLSAATPRHLKSGTLLSPDQEDNEIN 419
32 QY 417 HIOYFMQYKIDLPVTFREGANPGFHEAIGDIMALSVTPKHLYSLMLSTE-GSGYEYDIN 475
33 DB HIOYFMQYKIDLPVTFREGANPGFHEAIGDIMALSVTPKHLYSLMLSTE-GSGYEYDIN 475
34 QY 420 FLTLQALTYIGTTLPEFTMLKWRMAYFKGEIPADQMKKWMKREIYGVPEYPHDET 479
35 DB FLTLQALTYIGTTLPEFTMLKWRMAYFKGEIPADQMKKWMKREIYGVPEYPHDET 479
36 QY 476 FLTKALDKLTAIFPSTLIDQMKWRPFGDSITKENYNOBWSLRLKYGLOCPYPRBSGD 535
37 DB FLTKALDKLTAIFPSTLIDQMKWRPFGDSITKENYNOBWSLRLKYGLOCPYPRBSGD 535
38 QY 480 CDPASLIFVNSDYSFIRYTRTLVYQFQFQALQAAKHGEPRLHKCDISNTEAGOKLFNM 533
39 DB CDPASLIFVNSDYSFIRYTRTLVYQFQFQALQAAKHGEPRLHKCDISNTEAGOKLFNM 533
40 QY 536 FDPGSKFHVPAWVYRVFVSFIQFQFHALCRAAGHVGPHKCDIYOSKBAAGLADA 595
41 DB FDPGSKFHVPAWVYRVFVSFIQFQFHALCRAAGHVGPHKCDIYOSKBAAGLADA 595
42 QY 540 LRLKSESWTTLAENVGAKNNMNRPLNFEELFTWMLDKNR--NSFYGN-STDWSP 594
43 DB LRLKSESWTTLAENVGAKNNMNRPLNFEELFTWMLDKNR--NSFYGN-STDWSP 594
44 QY 596 MKLIGSKPWPAMKLTIGQPMMSASAMNFKPLTMTLVTEBRNGEITLGMPEYMAP 653
45 DB MKLIGSKPWPAMKLTIGQPMMSASAMNFKPLTMTLVTEBRNGEITLGMPEYMAP 653
46
47 RESULT 6
48 US-08-989-299-8
49 Sequence 8, Application US/08989299
50 Patent No. 6194556
51 GENERAL INFORMATION:
52 APPLICANT: Acton, Susan L.
53 APPLICANT: Robinson, Keith E.
54 TITLE OF INVENTION: ANGIOGENIN CONVERTING ENZYME HOMOLOG
55 TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
56 NUMBER OF SEQUENCES: 14
57 CORRESPONDENCE ADDRESS:
58 ADDRESSEE: FOLEY, HONG & ELIOT LLP
59 STREET: One Post Office Square
60 CITY: Boston
61 STATE: MA
62 COUNTRY: USA
63 ZIP: 02109-2170
64 COMPUTER READABLE FORM:
65 MEDIUM TYPE: Floppy disk
66 COMPUTER: IBM PC compatible
67 OPERATING SYSTEM: PC-DOS/MS-DOS
68 SOFTWARE: Patentin Release #1.0, Version #1.30
69 CURRENT APPLICATION DATA:

```


APPLICATION NUMBER: US/08/989,299
 FILING DATE: 11-DEC-1997
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Arnold E., Beth
 REGISTRATION NUMBER: 35,430
 REFERENCE/DOCKET NUMBER: MIA-025.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1312 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-989-299-8

Query Match 34.5%; Score 1334; DB 4; Length 1312;
 Best Local Similarity 42.6%; Pred. No. 5,8e-119;
 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

2 TIEOAKTFLEKNEHAEADLFYQSSLASWNTNTITEENVQNNNAGDKWSAFLEQSTL 61
 649 TDEAKADRYEEDRTAQLVLENYEAMNQNTNITIEGSKILEKSTEVSHHTLKYGTR 708
 62 AOMPLQELIOLTVKLOALQONGSVLSSEDSKRLNTILNTMSTIYSGKVCNPDNPQ 121
 709 AKTFVSNFQNSIKRIKILQNLDRALVPKELEBYNQILDMEYTSLSNICYING-- 766
 122 ECLLEPGNLEIMANSIDYNERLWMSRSEVQOLRPLYEEYVVKNAARAHNHYEDY 181
 767 TCMPEPLDTLNNMAKSRKEELLMAMKSWDRKVGRAILPEFPKYEVESKIKLNGYDA 826
 182 GDYMRGDEVNGVDGYDSRGQLIEDVEHTFEIKPLYEHLHAYVRKLMNAPS-YISP 240
 827 GDSWRSVYESDLE-----QDEKLYQELQPLYLNLHAYVRSLHRYGSEYINL 876
 241 ICGPLAHLLGDMGRFNTNLSLVPPGQKPNIDVTAMVDQAMDARIKEAEKFFVSV 300
 877 DGPPIAHLLGDMGRFNTNLSLVPPGQKPNIDVTAMVDQAMDARIKEAEKFFVSV 936
 301 GLPMNTOGFWMNSMLDPGNVOKAVCHPTAMDGLGK-DPRLIMCTKYTMDFLAHHEMG 359
 937 GLPVPPEFWMKSMLEKPTDGRREVYCHPSAMDYFNGKQFRLKQCTSVNMEDVIAHHEMG 996
 360 HIQDMAYAAOPFLIRNGANEGFHEAVGEIMSLSATPRLKLSIGLSPDFQEDNETEIN 419
 997 HIQYFMQYKDLVTFREGANPGFHEALIGDIMALSVTRKHLISLMLSTE-GSGYEYDIN 1055
 420 FLIKQALTYGLPTFTYMLEKRWNVFKEIPKQDMKRWMEKREIYGVVEPVPHDET 479
 1056 FLMKALDLKIAIFPSYLLIDQWRHVFDSITKENYDQEWMSLRKYQGLCPVPVRSQGD 1115
 480 GDPASLFHVSNDYSFIRYRTTLVQFOFQALCOAAKHEGPLKHKDISNSTAGOKLFNM 539
 1116 FDPGSEFHVANVPYIRYFISFIQFQHEALCRAAGHTGPLKHKDIIQSKAGKILADA 1175
 540 LRIKSEPTLLENVYGAKNANVRPLLYEPELFTWLKDNK--NSFVGM-STWSP 594
 1176 MKLGSKPVPKAMKILITQPPNNSASAMNVEFKLEMLVTEHRRIGETILGMEYIMAP 1233

RESULT 7
 US-08-989-299-9
 Sequence 9, Application US/08989299
 Patent No. 6194556
 GENERAL INFORMATION:
 APPLICANT: Acton, Susan L.
 APPLICANT: Robinson, Keith E.
 TITLE OF INVENTION: ANGIOGENESIS CONVERTING ENZYME HOMOLOG
 TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR

NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/989,299
 FILING DATE: 11-DEC-1997
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Arnold E., Beth
 REGISTRATION NUMBER: 35,430
 REFERENCE/DOCKET NUMBER: MIA-025.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1313 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-989-299-9

Query Match 33.9%; Score 1310; DB 4; Length 1313;
 Best Local Similarity 42.0%; Pred. No. 1.2e-116;
 Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps 7;

2 TIEOAKTFLEKNEHAEADLFYQSSLASWNTNTITEENVQNNNAGDKWSAFLEQSTL 61
 650 TDEKANKFVEYEDRTAKVLWNEYAEANWNTNITIEGSKILLQKKREVSNNHTLKYGTV 709
 62 AOMPLQELIOLTVKLOALQONGSVLSSEDSKRLNTILNTMSTIYSGKVCNPDNPQ 121
 710 AKTFVSNFQNSIKRIKILQNLDRALVPKELEBYNQILDMEYTSLSNICYING-- 767
 122 ECLLEPGNLEIMANSIDYNERLWMSRSEVQOLRPLYEEYVVKNAARAHNHYEDY 181
 768 TCMPEPLDTLNNMAKSRKEELLMAMKSWDRKVGRAILPEFPKYEVESKIKLNGYDA 826
 182 GDYMRGDEVNGVDGYDSRGQLIEDVEHTFEIKPLYEHLHAYVRKLMNAPS-YISP 240
 828 GDSWRSVYESDLE-----QDEKLYQELQPLYLNLHAYVRSLHRYGSEYINL 877
 241 ICGPLAHLLGDMGRFNTNLSLVPPGQKPNIDVTAMVDQAMDARIKEAEKFFVSV 300
 878 DGPPIAHLLGDMGRFNTNLSLVPPGQKPNIDVTAMVDQAMDARIKEAEKFFVSV 937
 301 GLPMNTOGFWMNSMLDPGNVOKAVCHPTAMDGLGK-DPRLIMCTKYTMDFLAHHEMG 359
 938 GLPVPPEFWMKSMLEKPTDGRREVYCHPSAMDYFNGKQFRLKQCTSVNMEDVIAHHEMG 997
 360 HIQDMAYAAOPFLIRNGANEGFHEAVGEIMSLSATPRLKLSIGLSPDFQEDNETEIN 419
 998 HIQYFMQYKDLVTFREGANPGFHEALIGDIMALSVTRKHLISLMLSTE-GSGYEYDIN 1056
 420 FLIKQALTYGLPTFTYMLEKRWNVFKEIPKQDMKRWMEKREIYGVVEPVPHDET 479
 1057 FLMKALDLKIAIFPSYLLIDQWRHVFDSITKENYDQEWMSLRKYQGLCPVPVRSQGD 1116
 480 GDPASLFHVSNDYSFIRYRTTLVQFOFQALCOAAKHEGPLKHKDISNSTAGOKLFNM 539
 1117 FDPGSEFHVANVPYIRYFISFIQFQHEALCRAAGHTGPLKHKDIIQSKAGKILADA 1176

QY 540 LRLKSEPTLALENVGAKNNVPLNTFEEPLTLKDKON--NSFVGM-STDMS 594
 Db 1177 MKLGSKOMPEAKKITGOPNMSASAIIMNYFKPLTEMLVTEENRRHGELTGMPEYTWTP 1234

RESULT 8

US-08-989-299-6
 ; Sequence 6, Application US/08989299
 ; Patent No. 6194556

GENERAL INFORMATION:

APPLICANT: Acton, Susan L.
 TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
 TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/989,299
 FILING DATE: 11-DEC-1997
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Arnold E., Beth
 REGISTRATION NUMBER: 35,430
 REFERENCE/DOCKET NUMBER: MIA-025.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 737 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-989-299-6

Query Match

Best Local Similarity 40.8%; Pred. No. 1.8e-114;
 Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;

QY 2 TIEBQAKTFLDKFNHAEADLFYQSSLASWNTNTITEE-----NVQNMNN--AGDKW 51
 Db 75 TDEAASRVEVEYDRSFOAVWNEAEANNNNTNTITEEASKILLQKMMQIANHLLTYGNW 134
 QY 52 SAKLEOSTLAOMTPILOEIONLVKLOLQONGSSVSEDSKRLNTIANTSTYST 111
 Db 135 -----ARRDVSNFQNMATSKRIKKVQDLQRAVLTPKKELEENQILLMETIYSV 184
 QY 112 GKVCNPDNPQECILLEGGLNIMANSIDYNERLWAMSESEVQKORPIYEEVVLKNE 171
 Db 185 ANVCQVDS--SCLDLPEPLNTLMATSKRYDELWMTSMRKGRAILTPFPKIVERTNK 242
 QY 172 MARANHVEDYDNGVYEVNGVDYDSRGQLTIEDVHTFEELKPIYEHLLHAYVRAKIM 231
 Db 243 AARLNGYDADDSRMSKMETPLE-----QDLERLFQELQPIYLIMHNAVGAALH 292
 QY 232 NAY-PSYISPTGLCPAHLIGDMGRFMTNLYSLTFPPGQKPNIDVTAMVDQANDQIRIF 290
 Db 293 RHYGAQHINTEGPIDPAHLIGNMMAQOTNSINIDYVAPPSASTMDATAMIKOGWTPRMF 352
 QY 291 KEAEFEVSVGLPMNTOGFEMNSMLTDPGNQKAVCPHAMDGKG--DEFLIMOTKYTMD 349

Db 353 EEADEFFISLGLPVPPEEFWNKSMLEKPTDREYVCHASANDFYNGKDFRIKQCTTYNME 412
 QY 350 DELTAHHEHGIQYDMAYAAOPELLRNGANEGFHEAVGEIMSLSATPRKLSIGLSPD 409
 Db 413 DLVYVHHEKHIOYFMQYKDLPVALREGANPGFHAIDYALSLASTPRKLSHINLSSE 472
 QY 410 FQEDNETEINFLKQALITVGTLPFTYMLEKRMVFGELIPKQDMKRMKREIYGV 469
 Db 473 -GGGYHDINFLMKMLDKIAFIPIPSYIVDEMRWVFDGSLTKENYNOEMWSLRKYQGL 531
 QY 470 VEPVPHDEYCDPASLFHVSNDYSPFRTYTRTLVQFOFQALCOAKHKGPLHCDJNS 529
 Db 532 CPAPRSQGFDPGAKFHPSSVPYIRFVSTIIOFHEALCKAAGHTPLTCDIYQS 591
 QY 530 TEAGOKLFMLRLKSEPTLALENVGAKNNVPLNTFEEPLTLKDKON--KNSFVG 587
 Db 592 KEAGKRLADAMKLGSKPPEAKMYITGPNMSASAMNYPPLMDWLTLENGRGELTG 651
 QY 588 W-STDMSPYADOS 599
 Db 652 WQYTWTPNSARS 664

RESULT 9

US-08-989-299-10
 ; Sequence 10, Application US/08989299
 ; Patent No. 6194556

GENERAL INFORMATION:

APPLICANT: Acton, Susan L.
 TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
 TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/989,299
 FILING DATE: 11-DEC-1997
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Arnold E., Beth
 REGISTRATION NUMBER: 35,430
 REFERENCE/DOCKET NUMBER: MIA-025.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1310 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-989-299-10

Query Match

Best Local Similarity 40.8%; Pred. No. 4.6e-114;
 Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;

QY 2 TIEBQAKTFLDKFNHAEADLFYQSSLASWNTNTITEE-----NVQNMNN--AGDKW 51
 Db 648 TDEAASRVEVEYDRSFOAVWNEAEANNNNTNTITEEASKILLQKMMQIANHLLTYGNW 707

```

OY 52 SAFLEOSTLAOMYPLDIOELNLYKLOALQONGSSVLSHDKSKRLNTLNTMSTYST 111
DB 708 -----ARRDVSNFQWATSKRIKIKVQDQRAVLVPEKLEENQJLLDMETYSV 757
OY 112 GKVCNPNPOECLELLEGLNEMANSLDYNERLWMSRSEVCKOLRPLXEYVYLKNE 171
DB 758 ANVCRRVGS--SCOLEPDLNLTAKTSKRYDELLMWTSMCKGRALLPFPKRYEPTNK 815
OY 172 MARANHYEDYDGYRGGYEVNGDYDYSRGOLEDYEHTEFEIKPLYEHLHAYRAKLM 231
DB 816 AARLNGYVADGDSWRSMYETPLE-----QDLERLFOELQPLYLMLHAYGALH 865
OY 232 NAY-PSYISPGCLPALLDGMGRFTNTLSLTPFGOKPNIDVDAMVDQAMQRIE 290
DB 866 RHGAQHINLEGPRLHLLGNMMAQWTSNTIDYVAPPSASTADATEAMKQGTPEBMF 925
OY 291 KEAEKFVSVGLPMTGFWENSMYLDPGVAVOKAVCHPTAMDGKG--DEFLMCTVTMD 349
DB 926 EADKPFISLGLLPVPEFNNKGMLEKPTDGREVYCHASAMDPYNGKDFIKOCTVYME 985
OY 350 DELTAHEMGHIYDAAVYAAOPFLRNGANEHGEHVAEGLMSLAATPKHLKSIGLSPD 409
DB 986 DLVAVHEMGHIYFMQKDLPAVALREGANPGEHAGLDVALSVSTPKHLHSINLSSB 1045
OY 410 FOEDNETEINFLKQALTYGTLPYTMLEKRMWYFKGPIPDQMKMKWEKKREIVGY 469
DB 1046 -GGGYEHNDINFLKMAKLDIAETPESTYLDWMRWVDSITIKENTNOEWMSLRLKYGL 1104
OY 470 VERVPHDEYCDPASLEHVSNDYSFTRYTRRLYOFQFQALCOAKHGEPLKCDISNS 529
DB 1105 CPAPPSQGDGDFGAFHLPSSVPIRYEVSFIIOFHEALCKAAGHCPGLHCTIYOS 1164
OY 530 TEAGOKLRLKRGKEPWTALLENVYGAKNNAVRLNYPEPLFTWLKDN--KNSFWG 587
DB 1165 KEKGKRLADMKRGYKSPWEAKVITGOFNMSASAMNYFKPLMDLLENGRHEGKIG 1224
OY 588 W-STDMSPYADQS 599
DB 1225 WFOYTWTPNSARS 1237

RESULT 10
US-08-989-299-11
Sequence 11, Application US/08989299
Patent No. 6194556
GENERAL INFORMATION:
APPLICANT: Acton, Susan L.
APPLICANT: Robinson, Keith E.
TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,299
FILING DATE: 11-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Arnold E., Beth
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MTA-025.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000

```

```

TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 615 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-989-299-11

Query Match
Best Local Similarity 36.5%; Pred. No. 2,2e-94;
Matches 218; Conservative 120; Mismatches 238; Indels 22; Gaps 9;

OY 4 EDAQKFFLDKFNHEADLFYQSSLSAWNTNTITEBNVONNNAAGDKWSAFLEQSTLAQ 63
DB 22 ETQAKYEYLENLKELAKRTNVETEAMAGNSITENEKKKEISAEILAKFEKVAASDT 81
OY 64 MYPLDIOELNLYKLOALQONGSSVLSHDKSKRLNTLNTMSTYSTGKVCNPNPOEC 123
DB 82 KFGWMSYQSEDLKQPKALTKIGYALPDDYAEILLDTLSMESNFAYKVCDDYKDSKC 141
OY 124 -LLEPGLNEMANSLDYNERLWMSRSEVCKOLRPLXEYVYLKNEMARANHYEDYG 182
DB 142 DLALPDEIEVYSKSRDHEELAYWREFYDKAGTAVRSOFERYVELNTRKAKLNFTSCA 201
OY 183 DYWRGDEVNGVDGYDYSRGOLEDYEHTEFEIKPLYEHLHAYRAKLMNAY-PSYISPT 241
DB 202 EAMLEDEY-----DQTFEQLDEDI---FADIRPLYOOIHGYAFRLRKHYGAVVSET 251
OY 242 GCLPALLDGMGRFTNTLSLTPFGOKPNIDVDAMVDQAMQRIEFAEKFEVSVG 301
DB 252 GPIPHILGNMMAQWSEIADIVSPERKPLVDYSAHEKQATYPIKAFQGGDDFTSMN 311
OY 302 LPMNQGWENSMYLDPGVAVOKAVCHPTAMDGK--GDFRIIMCTKVTMDDELTAHEMGH 360
DB 312 LTKLPQDWDMSLIERPFDGRDLVCHASAMDPYLDVIRIKQCTRVVQDOLFTVHHEIGH 371
OY 361 IOYDMAVAAOPFLRNGANEHGEHVAEGLMSLAATPKHLKSIGLSPDPOEDNETELNF 420
DB 372 IOYFLQYOHOPRYVTGTANPGFHEHGVYLSLSTVPHLEKIGLKL-DYVRDDEARINQ 430
OY 421 LKQALTYGTLPYTMLEKRMWYFKGPIPDQMKMKWEKKREIVGYVVERPHDEYTC 480
DB 431 LFTALDLYVLPFAFTMDKRWMSLFRGEVDKAMNCAFWLRIDREYSIGIEPPVRSKDF 490
OY 481 DEASLEHVSNDYSFTRYTRRLYOFQFQALC-QAAKH-----BGPLHKDISNTEAGQ 534
DB 491 DAPAKTHLSADVEYRKYRVVSFTIOFQFKSACIAGQYDPPDNVELPLDNCDIYGSARAGA 550
OY 535 KLFNMLRLKSEPTWALLENVYGAKNNAVRLNYPEPLFTWLKDN--KNSFWGMS 590
DB 551 AFHNMLSGASKPMPMDLLEAFNGERINSGKAIATFEPLRWLEAENIKNNVHIGWTT 608

RESULT 11
US-09-440-325A-1
Sequence 1, Application US/09440325A
Patent No. 6280994
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
TITLE OF INVENTION: Zaci: A Human Metalloenzyme
FILE REFERENCE: 98-79
CURRENT APPLICATION NUMBER: US/09/440,325A
CURRENT FILING DATE: 1999-11-15
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 694
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT

```

LOCATION: (1)...(694)
OTHER INFORMATION: Xaa - Any Amino Acid
US-09-440-325A-1

Query Match 25.6%; Score 989; DB 4; Length 694;
Best Local Similarity 37.9%; Pred. No. 3.2e-86;
Matches 221; Conservative 104; Mismatches 228; Indels 30; Gaps 15;

QY 4 EEOAKTEFLDKENHAEDELFGYSSLASWNTNTEENVQNNNAGDKMSAFLEQSTLAQ 63
DB 53 ETEETKLELQFYDGTGCVLTKMFKEMKNTVITNTRKNDPEMKMDER-SQFPIITGTAAH 111
QY 64 MYPLAEIONLTVKLOLALQNGSSVLSDEKSKRLNTLNTSTYTGKVCNPDNPEC 123
DB 112 LFKVQFKPDVNGMLSKLQNDKALSKDELREYNELLAXLEMTYSMAQVLNGBP--C 169
QY 124 LLEPGLEINANSIDYNERLWAMESREVGKOLRPLREEVYVKNEMARAHN-YEDYG 182
DB 170 LSLSEEL-EVMAISREKLELWAMOGMDAVGRQCTTFEHEVYELSKRAQLNGYXKDXG 228
QY 183 DVMRGDEVNGVDYDSRGOLIEDEHTEFEIKPLYEHLHAYRAKLNNAY-PSYISPI 241
DB 229 ALMHSYEDDTLE-----QDLERLFOELRPLYLNPHTYRRALHRRHGPEDIDR 278
QY 242 GCLPAHLIGD-MMGREWTNLYSLTFEFGKPNIDVTDAVDAQDAQRI-FKEAEKFTVS 299
DB 279 GPILAHLLGENTLAQSWNILDVLPFLKIKEDVTKIKKVOHMKREKIMLEAEETFFTY 338
QY 300 VG--LKNMTQGFENSMILDPGVNQAQCHPRAMDGK-GDRILMCTVTMDPFLTAH 356
DB 339 LGIALPPAPSEFKKMLMKPFDGVECHISAMNTYQDDDFRICKCAVTTEDPLSTH 398
QY 357 EMGHIOYDAAVYAAQPLLRNGANEGHEAVGEIMLSAATPKHLKSIGLSLDPQEDNT 416
DB 399 EMGHIOYDAAVYAAQPLLRNGANEGHEAVGEIMLSAATPKHLKSIGLSLDPQEDNT 453
QY 417 EINFLLKQALITVGLPFTYMLEKRWNVFKGEIPKDDMMKWKMEKR-EIYGVVPRPH 475
DB 454 EVNFMHIALEKIAFTPFGLTMDLFRMKVFDGTIMWDIYQEWNNLRRLKYQGLCPALH 513
QY 476 DENVCPASLEFHSNYSFI-RYVTRTYLQFQFQALCOAAKHEGLHCKDISNSTEAGQ 534
DB 514 SEEDPFGAKFHFSAGPYTRRFLSYLQFQFHEHLCKRASGMGLHQCDDIYNSIAK 573
QY 535 KLFNMLRLKSEFWTLALENVGAKANNVRLINTEPELFTWL 577
DB 574 LL--ALKIGSSKPPPEVLKMLTGESEVSTVWFMTYFKPLLTWL 614

RESULT 12
US-08-989-299-12
Sequence 12, Application US/08989299
Patent No. 6194556

GENERAL INFORMATION:
APPLICANT: Action, Susan L.
TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,299

FILING DATE: 11-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Arnold E., Beth
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MIA-025.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 907 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-989-299-12

Query Match 16.5%; Score 638.5; DB 4; Length 907;
Best Local Similarity 27.0%; Pred. No. 2.6e-52;
Matches 167; Conservative 132; Mismatches 280; Indels 39; Gaps 18;

QY 4 EEOAKTEFLDKENHAEDELFGYSSLASWNTNTEENVQNNNAGDKMSAFLEQSTLAQ 63
DB 178 EFKRLSLAGYEAIALIVLREVALLSGRYPNDASPILKALDEAENVLFVSTSMQAK 237
QY 64 MYPLAEIONLTVKLOLALQNGSSVLSDEKSKRLNTLNTSTYTGKVCNPDNPEC 123
DB 238 QFDMAVYDEKVMKOLGVSEFGMSALAPSRFADYSQAQALNDSKSTICDKDVPFC 297
QY 124 LLEPGLEINANSIDYNERLWAMESREVGKOLRPLREEVYVKNEMARAHNEDYGD 183
DB 298 ALQKIDMOSITRNEKQASRLQHLWVSYTAIAKS-KSYNNITITSNKGAKNGFANGA 356
QY 184 YWRGDYEVNG-VGDYDSRGOLIEDEHTEFEIKPLYEHLHAYRAKLNNAY-PSYISPI 240
DB 357 MWRSAFDMSSKYHKAFF--DLNKQIDKISTIQPFYQDLHAYRRQLAGIYSNPVLSK 413
QY 241 IGCPLAHLIGDMKGFNTNLSLYVFPQKPNIDVTAMVD---QAMDQRIKFAKE 296
DB 414 DGPILAHLLGSLDGDMSAHYEQKPFEEES--ETPEAMLSAFTONTTKMFTVYARY 471
QY 297 FVSVGLPMMQGFENSMILDPGVNQAQCHP-TAMD-L-GKDRILMCTVTMDPFLTA 354
DB 472 EKSAGFPHLPKSYWSSIFARVWS-KMICHPAALDRAVNDPRVYKCAQLGSPDEQA 530
QY 355 HHEMGHIQYDAAVYAAQPLLRNGANEGHEAVGEIMLSAATPKHLKSIGLSLDPQEDN 414
DB 531 HSLVQTYQYLYKQDSLLFREQASPVITDAIANAFHLSNPHYLSQKLVSEHDIK 590
QY 415 ETE-INFLKQALITVGLPFTYMLEKRWNVFKGEIPKDDMMKWKMEKREIYGVVEPV 473
DB 591 DSVIINKLKRESLEFETKLPITLADNNRYELFDQTVKKNLNDRWMEIRNKYGVSPQ 650
QY 474 PHDETYCPASLEH--VSNDYSFIRYRTL---YQFQDALCOAA---KHGRLHCK 524
DB 651 PYNTSNLD--ALIHNSVSQVHS---PATRTLSYVLKQIILKALCORLFWLSG---C 701
QY 525 DISNSTEAGQKLFNMLRLKSEFWTLALENVGAKANNVRLINTEPELFTWLKQDNK-- 582
DB 702 ILSEDTT-EKLRETKLGSSITWLKALEMTISGKELDQAPLEYEYELINLNRNTNEID 759
QY 583 NSFVGMSTDMSPYADQSI 600
DB 760 QVVGWMDGEGTPTFYVEI 777

RESULT 13
US-08-645-193B-15
Sequence 15, Application US/08645193B
Patent No. 5962253
GENERAL INFORMATION:
APPLICANT: Kupke, Thomas

APPLICANT: Gotsz, Friedrich
 APPLICANT: Kempster, Christoph
 APPLICANT: Jung, Gunther
 TITLE OF INVENTION: Oxidative Decarboxylation of Peptides
 TITLE OF INVENTION: Catalyzed by Flavoprotein Epd
 NUMBER OF SEQUENCES: 70
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 STREET: 1100 New York Avenue, Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/645,193B
 FILING DATE: 13-MAY-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Esmond, Robert W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 0652,1540000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2540
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 990 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: not relevant
 MOLECULE TYPE: protein
 US-08-645-193B-15

Query Match 3.1%; Score 120; DB 2; Length 990;
 Best Local Similarity 18.2%; Pred. No. 0.023;
 Matches 150; Conservative 116; Mismatches 246; Indels 310; Gaps 40;

```

QY 2 TIEBOATFLDKFNH-----ADL--FIOSSIASWNYMTTEENVQMNNAADKW 51
DB 343 TINIYHEFDKGFQOLVNLKQLSDINGFGYPRKDSYFSGNNI----- 387
QY 52 SAKLKQOSTLAQMPLOEIONLVKLOALQONSGSVSEKSRMTIINTASTYST 111
DB 388 -AFLKER-----YLLAIONNSHIEITENDVNLK-NNTYSKINA- 425
QY 112 GKVCNPDNPQECLELPGINETMANSLDYNERLWAMESRSEVGRQLPLYEEYVVKNE 171
DB 426 -----PVSTEL-----YSELYF-----GNSING-YEDRAVISPI 453
QY 172 MARAHYEDYGDYWRGDEVNGVDYDYSRGQLIEDVHTPEELKPLYEHLAAVRAKLM 231
DB 454 LGSFNAGATFRF-TGNFNK-----KKNOLOKEIVH-----HYNNMTMNDL 495
QY 232 NAYPSIYISIGCLPAHLIGDMGRWTVLSLVPEFGKPNIDYTDAMVDQAPAGRIFK 291
DB 496 ELSQUNEAPLNSRNVILNN--NRIYNTCLMNLN--KSIDINDIFIGATFKKLYIS 550
QY 292 E--AEKFFVSGLPNMTQG-----FWNSMLTDRN-----VOKA 324
DB 551 EKHDSRIYFVNSMNTYFSGSELYKFLREISFEKTKIOTPTTEGIDSLFCPRILYKNI 610
QY 325 VCHPAMLDGKDRILIMCTVTMDLFLAHHEMGHIQYMAVA----- 368
DB 611 ILKPATWKINSEMESE--TENMLNRPATI-REKWHIPKVIYLAFGNRLMLLMDKHI 666
QY 369 -----AOPFLRNGANEGHEAVGEIMS-----LSAATPK-----HK 401
DB 667 IILKELKKHGRIRILSLFINESNNEMLEIVPLKYKTSLKDOSFLIRNRNKHNNK 726
  
```

```

QY 402 ----SIGLSPDQEDNETENPLKQALITVGLPTLYLKWMMVEKGEIPKQWKK 458
DB 727 DWSIHSLIPKTYQD-----NFIQDYLLPFITELKYNNEINFFYIKFKED---EDFK- 777
QY 459 WMEKREIYGVBPVHDETYCDPASLFVNSDYSFIRY-----TITLYQFOFQALC 512
DB 778 -LNLRE-----DEDY-----SQYIFIKMWDYCLINSELYDYSIVYVP 817
QY 513 QAAKHGRLHKDISSTAGOKLFTMLRLKSEFWTLALENVGAKMNVRLPLNTEEP 572
DB 818 EYTRIGEPVVIDIENFMYDSL--SINIIOSE-FKIPKEPIYA---ISIDFLDLE- 870
QY 573 LFTWLKDQKNSFVGWSTWSPYDOSIKVRISLSKALGDAYMWNENMYLFRSSVAYA 632
DB 871 -----INKSKEELLNNA--EDLYRSD----- 892
QY 633 MROYFLAKNQMILGDEADVANKLPRISEFVTAPKKNVSDIIP----- 678
DB 893 IREY-----KNLLAKL-----TJPKNDYELIKKEFPNLEFLNKI 928
QY 679 -----REVEKAIKMSRSRINDAF---RLDNSLEFLGTOP 711
DB 929 SIENLKKTKLOKSLYTSRSRIIGSFIMKCN---RIFGINP 966
  
```

RESULT 14
 US-08-630-916A-46
 Sequence 46, Application US/08630916A
 Patent No. 6011137
 GENERAL INFORMATION:
 APPLICANT: Pirozzi, Gregorio
 APPLICANT: Kay, Brian K.
 TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
 NUMBER OF SEQUENCES: 124
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/630,916A
 FILING DATE: 03-APR-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MISROCK, S. LESLIE
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 1101-203
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 896-8864/9741
 INFORMATION FOR SEQ ID NO: 46:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 683 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 US-08-630-916A-46

Query Match 3.0%; Score 116; DB 3; Length 683;
 Best Local Similarity 18.3%; Pred. No. 0.031;
 Matches 141; Conservative 107; Mismatches 265; Indels 258; Gaps 36;


```

QY 447 KGIIPKDOMKKMKWEKREIVGVEPVPHDEYCDPASLFHSNDYSFIRY-----TR 500
Db 769 KED---EDFK--LRLIRE-----DEDY-----SOIYSFIKNWDYCLNS 804
QY 501 TLVOFOFOALCOAKHGEPLHKCDISNSTEAGOKLFNMLRLKSEPMFLALENVGAKN 560
Db 805 ELIYISIVDYVPEYKRGPHVEDIENF--MTDSLDSINIIQSE--FKIPKEFYA--- 859
QY 561 MNVRPLNTEPEPLFTWLKDONKNSFWGSTDWSPYADOSIKYRISLSKALGKAYEMNDN 620
Db 860 ISIDFLDYLE-----INSEKEEILLNNA--EDLYRSND-- 892
QY 621 EMILFRSSVAYAMROYFLAKYKNOMILFGEEDVRYANLKPRISNFVTAPKAVSDIIP-- 678
Db 893 -----IREY-----KNLAKL-----TPKNKDYELIKKE 916
QY 679 -----RTEVEKAIKMSRSRINDAF--RLNDSLEFLGIOP 711
Db 917 FPNLHEFLFKISILENMLKKTLOKSLYTSRSRIIGSFIMRCN-----RIFGINP 966

RESULT 16
US-08-466-961A-20
: Sequence 20, Application US/08466961A
: Patent No. 5843709
: GENERAL INFORMATION:
: APPLICANT: Entlian, Karl-Dieter
: APPLICANT: Gtz, Friedrich
: APPLICANT: Schnell, No. 5843709bert
: APPLICANT: Augustin, Johannes
: APPLICANT: Engelke, Gerhard
: APPLICANT: Rosenstein, Ralf
: APPLICANT: Kaletta, Corlina
: APPLICANT: Klein, Cora
: APPLICANT: Wieland, Bernd
: APPLICANT: Kupke, Thomas
: APPLICANT: Jung, G nther
: APPLICANT: Kellner, Roland
: TITLE OF INVENTION: Biosynthetic Process for the Preparation of
: NUMBER OF INVENTIONS: 42
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox
: STREET: 1100 New York Avenue, NW
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/466,961A
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/392,625
: FILING DATE: 22-FEB-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/876,791
: FILING DATE: 30-APR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/784,234
: FILING DATE: 31-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Esmond, Robert W.
: REGISTRATION NUMBER: 32,893
: REFERENCE/DOCKET NUMBER: 0652.0980004
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540

```

```

: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 990 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
US-08-466-961A-20

Query Match      3.0%  Score 115; DB 2; Length 990;
Best Local Similarity 17.9%  Pred. No. 0.071;
Matches 149; Conservative 116; Mismatches 235; Indels 334; Gaps 41;

2 TIEQAKTFLDKFNHE-----AEDL--FYOSLSAMWNTNTITEENQNNMAGDKW 51
Db 343 TIRNHEFFMDKYGEGELVNLKOLLSDINGRCYKPKDSYSSNNI----- 387
QY 52 SAFLKEOSTLAOMPLQIOIWLTKLOLQLOAGSSVLSDESKRLNTILNTSTYST 111
Db 388 -AFLKEK-----YLAIOIWNHSHIETENDVKMLEK--NNTVSKINA-- 425
QY 112 GKVCNPDNPQCLLLEPLNEIMANSIDYNERLWAMESWRSEVGKOLRPLYEEVYLKNE 171
Db 426 -----PVSTEI-----YSRIY-----GNSIKG--YEDPAVISP 453
QY 172 MARANHEEDYGDYKRGDYEVNGVDGYDSRGQIIEDEHFEETKPLYEHLHAYRAKIM 231
Db 454 LGSEFAGATFGRP--TGNEFNK-----KKNQLOKEIYH-----HYNNMNMENDL 495
QY 232 NAYPSYISPICLRAHLILGDMKGRFNTNLSLYVPFGOKPRIDYTDAMQADQARIK 291
Db 496 EISOLNEAPLNSRVNINLN--NRIYNTCLNLNP--KSDIDNIDIFIGATFNKLYLS 550
QY 292 E--AEKEFVSGLPNNMTOG-----FWENSMITDGN-----VOKA 324
Db 551 EKHDSRIYFVSNSMNFNEFGSELYKPLREISFEKTKIQPTTEGIDSLPPCPRIITYKI 610
QY 325 VCHPTAMDLGK-----GDFRIIMCTRYTMDDEFLTAH 355
Db 611 ILKPATWKINSEMESETEMNLRPATIRKWHIPKXIVITARGDNRLIL-----NLNDK 663
QY 356 H-----EMGHIOYDMAVAAPFLIRNGANGCHEANGELMS-----LSAAT 396
Db 664 HLIILKEELKKHGRIR-----ILSEFINESNNERMLEVTYPLPKYKTSLEKESFII 713
QY 397 PK-----HLK-----SIGLSPDOENETEINFLKQALTIYGTLPFTYMLEKRWMPV 446
Db 714 PKNNKHPNNLXDMFSLHLSIPKYQD-----NFIDYLLPPTTELKVNNEFINKPFYIKF 768
QY 447 KGIIPKDOMKKMKWEKREIVGVEPVPHDEYCDPASLFHSNDYSFIRY-----TR 500
Db 769 KED---EDFK--LRLIRE-----DEDY-----SOIYSFIKNWDYCLNS 804
QY 501 TLVOFOFOALCOAKHGEPLHKCDISNSTEAGOKLFNMLRLKSEPMFLALENVGAKN 560
Db 805 ELIYISIVDYVPEYKRGPHVEDIENF--MTDSLDSINIIQSE--FKIPKEFYA--- 859
QY 561 MNVRPLNTEPEPLFTWLKDONKNSFWGSTDWSPYADOSIKYRISLSKALGKAYEMNDN 620
Db 860 ISIDFLDYLE-----INSEKEEILLNNA--EDLYRSND-- 892
QY 621 EMILFRSSVAYAMROYFLAKYKNOMILFGEEDVRYANLKPRISNFVTAPKAVSDIIP-- 678
Db 893 -----IREY-----KNLAKL-----TPKNKDYELIKKE 916
QY 679 -----RTEVEKAIKMSRSRINDAF--RLNDSLEFLGIOP 711
Db 917 FPNLHEFLFKISILENMLKKTLOKSLYTSRSRIIGSFIMRCN-----RIFGINP 966

RESULT 17
US-08-844-059-2
: Sequence 2, Application US/08844059
: Patent No. 6001601
: GENERAL INFORMATION:

```

APPLICANT: Lawlor, Elizabeth
 TITLE OF INVENTION: No. 6001601el Compounds
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham Corporation
 STREET: 709 Swedeland Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19046
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/844,059
 FILING DATE: 18-APR-1997
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 9607999.1
 FILING DATE: 16-APR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Gimmil, Edward R.
 REGISTRATION NUMBER: 38,891
 REFERENCE/DOCKET NUMBER: P31456
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090
 TELEEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 665 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-844-059-2
 Query Match
 Best Local Similarity 19.2%; Score 113.5; DB 3; Length 665;
 Matches 135; Conservative 99; Mismatches 231; Indels 239; Gaps 33;
 QY 125 LLEPGINELMANSIDYNERIAMESEKSEVQKQRLPYEEVYVILKEMARANHEDYGY 184
 DB 85 LLDISYDKFRTDYEHEKVA-----QVFERLLAQDDIY--LGEY 123
 QY 185 WRGDYEVNGVDGYDYSRGOLIE-----DVHTFEERIKPLYEHLNAVY 226
 DB 124 -SGWTSVS--DEEFFESQLAIEYFRDAGNVTGGIAPSGHEHVMWSESEYFLRLSKYQDR 180
 QY 227 RAKIMANVPSYISPIGCLPAHL-----LGDMMGRFTNLSLVVFGOKPNIDY-TDA 278
 DB 181 LVFEFFAHNPEFIRPDGLNMLNFIPEGLDLAVSRFTFWGVVPSPMKHVVYVWIDA 240
 QY 279 MV-----DQAWDA--QRTFKAEKF-----FVSGIPLN--MQ 307
 DB 241 LLNYATATGTAODEHNGFNDFKFWNGTFVHWGKDLFRHSLIYPIILLMLDVKLPRLIAH 300
 QY 308 GFV--ENSMLTDPGNVQKAVCPRTAMDGKGDGFRILMCTKVIMDDFLTAHHEMGIOYD 364
 DB 301 GNVVMDGKMSKSGNVVYREMLVERGYGLDPLRYILMRNLPVSDGTFPRDEYVGLINTE 360
 QY 365 MAYAAPFLTR-----NGANEGFHEAVGEIMSLSATPKHLKSGILSPDQEDNE 415
 DB 361 LANDIGLMLNRYTSMINKYFDGQIPAYVEGVTEFDHVALEVAE--KSIDA---DFRTIME 414
 QY 416 TEINFLKQALLTYGIL--PTTYMLEKRWAMFKGEIPRQDQ--MKKMKMKRRIYGV 469
 DB 415 A-VDY--PRALIEVWTLISRTNKYIDETAPWIDKDEALDQLASVSHQASIRVVAHL 471
 QY 470 VEPHPDEYCDPASLEHVSNDYSFIRYVYRTLYQFOFQALQAAKHGEGPLKHCDSINS 529

DB 472 IEP-----FMNSTRNAV-----483
 QY 530 TEAGCKLFNMLRCKSEFWTLAENV-----VGAKNMVPLNFEELFTWLKD 579
 DB 484 -----LTQIDLEE--VSSLNLSLADFPADVIVVANGTPIFPLNNEEEL-AVKE 531
 QY 580 ONKNSFYGWSTDMSP-----YAD-QSIRKIRIS-----LKSALGDKAYEW-- 617
 DB 532 QMEGNKPAVEKEWNPDEVELKLNKDEIKFEDEKVEIRVAVEKSVKESGDKILOFRID 591
 QY 618 -NDNEMYLFRSSVAVAROVFLKYNQMLIFGEEDVAVANLKPRISEFVTAPKNVSDI 676
 DB 592 AGDEGRQILSGIA-----KTY--PNQDELVGKVVQIVANLKR-----628
 QY 677 IPRTVEKAIMSRSRINDAERL--NDNSLEFLGIQPTLGPVN 717
 DB 629 -----KMKKVVYSGQMLLSAHDGKTLTLTVDPVAV--PN 660
 RESULT 18
 US-09-431-202-2
 ; Sequence 2, Application US/09431202
 ; Patent No. 6294175
 ; GENERAL INFORMATION:
 ; APPLICANT: Lawlor, Elizabeth
 ; TITLE OF INVENTION: No. 6294175el Compounds
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SmithKline Beecham Corporation
 ; STREET: 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19046
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/431,202
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/844,059
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gimmil, Edward R.
 ; REGISTRATION NUMBER: 38,891
 ; REFERENCE/DOCKET NUMBER: P31456
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-270-4478
 ; TELEFAX: 610-270-5090
 ; TELEEX:
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 665 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-431-202-2
 Query Match
 Best Local Similarity 19.2%; Score 113.5; DB 4; Length 665;
 Matches 135; Conservative 99; Mismatches 231; Indels 239; Gaps 33;
 QY 125 LLEPGINELMANSIDYNERIAMESEKSEVQKQRLPYEEVYVILKEMARANHEDYGY 184
 DB 85 LLDISYDKFRTDYEHEKVA-----QVFERLLAQDDIY--LGEY 123
 QY 185 WRGDYEVNGVDGYDYSRGOLIE-----DVHTFEERIKPLYEHLNAVY 226

```

Db 124 -SGWTSVS--DEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSESEYFLRLSKYQDR 180
QY 227 RAKIMNAPSYISPIGCLPAHL-----LGDMMGRFNTLYSLVPGOKPNIDY-TDA 278
Db 181 LVEFFKAHEFTIPDGRNENLRNIEBGDLAVSRITTFWGPVPSNPKHVVYVWIDA 240
QY 279 MV-----DQAWDA---ORITKEAKR-----FVSGLPN--MTQ 307
Db 241 LLNATATAGYADEHGNDFKFMNGTVFHMWKGDLRFHSIYPIILLMMLDVKLPLDLIAH 300
QY 308 GFW--ENSMLTDGPNQKAVCHPTAMDIGKDRILLMCTKVTMDLTLAHHEMGIQYD 364
Db 301 GMFWKQKMSKSGKNVYPPMLVERGLPLRYLLMKNLPVSGDGTFFPEDYVGRINYE 360
QY 365 MAYAAQPELRL-----NGANEGRHEAVGEIMSLSAATPKHLKSIGLSPDFOEDNE 415
Db 361 LANDLGNLNTVSMINKYEDGQIPAYVEGTEFDHVAEAE--KSLA---DFHIME 414
QY 416 TEINELLKQALTYIGTL--PFTYMLKRWMMYKGLIPDOW--KKKWMKREITGV 469
Db 415 A-VUY--PRALEAVWTLISRTNKYIDETAPWVLDKDALRDLQASVSHQASIRVVAHL 471
QY 470 VEPVHDEYCDPASLFEVSNDSYFIRYTRITYQFOFQALCOAAKHEGRLKCDLSNS 529
Db 472 IEP-----FMMETSRAY-----483
QY 530 TEAGQKLFNMLRLKSGSEPTLALENV-----VGARNMVRPLNTFEPLTWLKD 579
Db 484 -----LTQIGLEB--VSSLENTSLADPADVYVAKGTPLFPLRNLNEEI-AVIKE 531
QY 580 QNKSFWGSTMDSPP-----YAD-QSIVKRIS-----LKSALGDAYEY---617
Db 532 QMGNKRAVEKENPDEVELKKNDEKFEDEPKVEIRVAEVEVSVEGSDTLQFRID 591
QY 618 -NDNEMTLFSSVAYAMROYFLKVKMOMILFGEDEVYANLKPRISENFVTAPKRVSDI 676
Db 592 AGGEDRQLISGIA---KYY---PNDQELVGKKQVYANLKR-----628
QY 677 IPRTEVEKATIRMSRSRINDAFRL--NDNSLEFLGIQPIFGPPN 717
Db 629 -----KMMKKYVSGMILSAEHDKLTLLTVDPAV--PN 660

RESULT 19
US-08-491-357-2
; Sequence 2, Application US/08491357
; Patent No. 5716782
GENERAL INFORMATION:
APPLICANT: Golemis, Erica A.
APPLICANT: Law, Susan
TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING A SIGNAL
TITLE OF INVENTION: MEDIATOR PROTEIN THAT INDUCES CELLULAR MORPHOLOGICAL
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dann, Dorfman, Herrell and Skillman
STREET: 1601 Market Street Suite 720
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/491,357
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.

```

```

REGISTRATION NUMBER: 36,252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-491-357-2

Query Match
Best Local Similarity 20.5%; Score 104; DB 1; Length 834;
Matches 88; Conservativity 59; Mismatches 142; Indels 140; Gaps 19;

QY 38 EENVQNMNAGDKWSAFKE-----OSTLAQMTPIQE-IQNTVTKLOALQOQSSVYS 91
Db 442 ERIINERTAVDKVELFELKELYHFKAVANAAACLPILLHMKRREIQRV-EDSHQILS 500
QY 92 E-----DKSKRLNTLNT-----104
Db 501 QTSHDINECSWSLNLIAINKPQNKCDLDRFYVAKTYVDDAKQLTTINTINAEALFRPG 560
QY 105 -----NSTIY-----STGKVCNP-DNPOEC--LLEBGLNEIMA--NSIDY 140
Db 561 PGLHLKNGPESIMNSTEYPRHGSOGQLLHPGDHKAQAHNKLPLGLSKQAPDSSSDG 620
QY 141 NERLWAME-----SWRSVGRQLRPIYEEYVVLNENARAHYDYDGYRGRDYEVGV 194
Db 621 SERSMDDVDYHLDGKEFEFEQOKELLEKENIMQNKQLEHNO-----665
QY 195 DQYDYSKGLIDVHTHEEIRPLYEHLAVYRAKIMNAPSYISPIGCLPAHLIGDWG 254
Db 666 -----LSQFOLLQ-----EITKPVENDISKW--KPSGLPTNNGVSAQROLLCFYID 713
QY 255 REMTNYSLTVFPGKPNIDVDAMVDAMDAQ--RIKREKEF-----FVSGLPNM 305
Db 714 QCEITHRISL-----LNAIDALFSCVSSAQPRIYVANSKFEVILSAHRLVIG-DTL 763
QY 306 TQGFWENSLTDPGVOKAVCHPTAMDIGKDRILLMCTKVTMDL--TLAHHEMGIQY 363
Db 764 TRQVTAQDIRNKVNMSSNQLCEQLK-----TIVATKAAHLHPSTIALQEKVHOYT 815
QY 364 DMAYAAQEP 372
Db 816 DISRNAQLE 824

RESULT 20
US-08-968-633-2
; Sequence 2, Application US/08968633
; Patent No. 6100384
GENERAL INFORMATION:
APPLICANT: Golemis, Erica A.
APPLICANT: Law, Susan
TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING A SIGNAL
TITLE OF INVENTION: MEDIATOR PROTEIN THAT INDUCES CELLULAR MORPHOLOGICAL
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dann, Dorfman, Herrell and Skillman
STREET: 1601 Market Street Suite 720
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,633
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/491,357
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.
REGISTRATION NUMBER: 36,252
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
STRANDEDNESS: No. 6100384 Relevant
TOPOLOGY: No. 6100384 Relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-968-633-2

Query Match 2.7%; Score 104; DB 3; Length 834;

Best Local Similarity 20.5%; Pred. No. 0.61; Mismatches 142; Indels 140; Gaps 19;

Db 442 ERHINERTRAVDVELFELKYLHFGVAVANACLPDLITLHNMKRELQRY-EDSHQILS 500
QY 92 E-----D-----DKRLNTIINT----- 104
Db 501 QTSHDINECSWSLITLAINPKKCDLDRFVAVATVPDDAKOLTTTINTNAEALFRPG 560
QY 105 -----MSTY-----STGRVCP-DNPOEC--LLEPGLINEIMA---NSLDY 140
Db 561 PGSLHLKNGPESTINSTEYFHGSGOGLHFGDHKAQAHNKALPGLSKQAPDCSSDQ 620
QY 141 NERLAME-----SWREVGKOLRPLYEYVVLKNEANAHYEDYGDYWRGDEYNGV 194
Db 621 SERGMDDYDVHLQGEFEERQOKELLEKENIMKOKMQLHHQ----- 665
QY 195 DGYVYSGQLIEDVEHFEELKPLYEHLHAYVRAKLNNAYPSYISPGICLPAHLGDMWG 254
Db 666 -----LSOFQLEQ-----ETTKPEVDISKW---KPSQLPTNSGVSAQDRQLCEYD 713
QY 255 RFTWNLVSLVFPFGKPNIDVTDAVDAQ--RIKFAEKF-----FVSVGLPNM 305
Db 714 QCEPHFTSL-----LNAIDLFCVSSAOPRIFVAHSKFLVLSAKLVFIG-DTL 763
QY 306 TQGWENSMILTPGCVKAVACHPTAMDLGKDFRILMCTVYIMDF--LTAHHEMGHIQY 363
Db 764 TRQVTAODIRNKVANNSSNOLEQDK-----TIVATKMAALHPSTTALQEMVHOYT 815
QY 364 DMAYAAQPF 372
Db 816 DISRNAQLF 824

Search completed: May 26, 2003, 17:58:18
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 17:54:24 ; Search time 39 Seconds

(without alignments)
1830.880 Million cell updates/sec

Title: US-09-978-385-2_COPY_19_738

Sequence: 1 STIEEQAKFTLDFKFNHEAD.....DNSLRFGLGTPPLGPPNDPP 720

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	3869	100.0	805	9	US-09-978-385-2
2	3869	100.0	805	9	US-10-158-847-142
3	3865	99.9	805	9	US-10-114-893-86
4	3620	93.6	681	9	US-09-968-384-25
5	3620	93.6	681	9	US-10-138-847-140
6	3620	93.6	711	9	US-09-969-384-13
7	3620	93.6	711	9	US-10-158-847-138
8	3251	84.0	805	9	US-09-978-385-6
9	3233	83.6	805	9	US-09-978-385-9
10	2897	74.9	555	9	US-10-028-072-72
11	2897	74.9	555	9	US-10-121-049-72
12	2897	74.9	555	9	US-10-123-904-72
13	2897	74.9	555	9	US-10-140-470-72
14	2897	74.9	555	9	US-10-175-746-72
15	2897	74.9	555	9	US-10-176-918-72
16	2897	74.9	555	9	US-10-176-921-72
17	2897	74.9	555	9	US-10-137-865-72
18	2897	74.9	555	9	US-10-140-474-72
19	2897	74.9	555	9	US-10-142-431-72

20	2897	74.9	555	9	US-10-143-114-72	Sequence 72, Appl
21	2897	74.9	555	9	US-10-140-002-72	Sequence 72, Appl
22	2897	74.9	555	9	US-10-142-419-72	Sequence 72, Appl
23	2897	74.9	555	9	US-10-123-262-72	Sequence 72, Appl
24	2897	74.9	555	9	US-10-142-423-72	Sequence 72, Appl
25	2897	74.9	555	9	US-10-121-050-72	Sequence 72, Appl
26	2897	74.9	555	9	US-10-141-755-72	Sequence 72, Appl
27	2897	74.9	555	9	US-10-143-032-72	Sequence 72, Appl
28	2897	74.9	555	9	US-10-123-108-72	Sequence 72, Appl
29	2897	74.9	555	9	US-10-123-236-72	Sequence 72, Appl
30	2897	74.9	555	9	US-10-123-261-72	Sequence 72, Appl
31	2897	74.9	555	9	US-10-140-921-72	Sequence 72, Appl
32	2897	74.9	555	9	US-10-140-928-72	Sequence 72, Appl
33	2897	74.9	555	9	US-10-121-045-72	Sequence 72, Appl
34	2897	74.9	555	9	US-10-123-292-72	Sequence 72, Appl
35	2897	74.9	555	9	US-10-123-903-72	Sequence 72, Appl
36	2897	74.9	555	9	US-10-124-819-72	Sequence 72, Appl
37	2897	74.9	555	9	US-10-124-822-72	Sequence 72, Appl
38	2897	74.9	555	9	US-10-140-925-72	Sequence 72, Appl
39	2897	74.9	555	9	US-10-160-498-72	Sequence 72, Appl
40	2897	74.9	555	9	US-10-121-041-72	Sequence 72, Appl
41	2897	74.9	555	9	US-10-121-043-72	Sequence 72, Appl
42	2897	74.9	555	9	US-10-121-047-72	Sequence 72, Appl
43	2897	74.9	555	9	US-10-123-215-72	Sequence 72, Appl
44	2897	74.9	555	9	US-10-123-902-72	Sequence 72, Appl
45	2897	74.9	555	9	US-10-123-908-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1
US-09-978-385-2
Sequence 2, Application US/09978385
Patent No. US2002017221A1
GENERAL INFORMATION:
APPLICANT: Piddington, Christopher S.
APPLICANT: Petrie, Charles
APPLICANT: Shoemaker, Kimberly E.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: ZINC2: A HUMAN METALLOENZYME
FILE REFERENCE: 99-24C1
CURRENT APPLICATION NUMBER: US/09/978,385
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 60/133,952
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: 60/151,181
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: 09/563,516
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 805
TYPE: PRT
ORGANISM: Homo sapiens

Query Match 100.0%; Score 3869; DB 9; Length 805;
Best Local Similarity 100.0%; Pred. No. 1.9e-312;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 STIEEQAKFTLDFKFNHEADLFYSSLSASWNTNTTEENVOMNNAAGKMSAFLKEOST 60
19 STIEEQAKFTLDFKFNHEADLFYSSLSASWNTNTTEENVOMNNAAGKMSAFLKEOST 78
61 LAQMYPLQETONTLVKIQDALQONGSSVLSSEKSKRLNTIINTMSTISTGKVCNPNP 120
79 LAQMYPLQETONTLVKIQDALQONGSSVLSSEKSKRLNTIINTMSTISTGKVCNPNP 138
121 QECILLEPGINEIMANSIDYNERITMAMESRSRSEVGKOLPIVEEYVILKEMARAHYED 180
139 QECILLEPGINEIMANSIDYNERITMAMESRSRSEVGKOLPIVEEYVILKEMARAHYED 198

QY 181 YGDYWRGDEYVNGVGYDYSRGQLIEDYEHTEFEIKPLYEHLHAYVRAKLMAAYPSYISP 240
DB 199 YGDYWRGDEYVNGVGYDYSRGQLIEDYEHTEFEIKPLYEHLHAYVRAKLMAAYPSYISP 258
QY 241 IGCPLAHLGDMGGRFTNLYSLTPFGOKPNDIVTDAMVDQAMDQORIFKEAEKFFVS 300
DB 259 IGCPLAHLGDMGGRFTNLYSLTPFGOKPNDIVTDAMVDQAMDQORIFKEAEKFFVS 318
QY 301 GLPNTGFWENSMLTDPGNVOKAVCHPTAMDLCGDRILMCTRYTMDDFLTAHHEMGH 360
DB 319 GLPNTGFWENSMLTDPGNVOKAVCHPTAMDLCGDRILMCTRYTMDDFLTAHHEMGH 378
QY 361 IQYDMAYAAOPFLLRNGANGFHEAVGEIMSLAATPKHLKISGLSPDFQEDNETEINF 420
DB 379 IQYDMAYAAOPFLLRNGANGFHEAVGEIMSLAATPKHLKISGLSPDFQEDNETEINF 438
QY 421 LKQALITVGLPTFTYMLEKRWNVFEGEIPKQDMKKWEMKREITGVVEVPDDETYC 480
DB 439 LKQALITVGLPTFTYMLEKRWNVFEGEIPKQDMKKWEMKREITGVVEVPDDETYC 498
QY 481 DPASLFHVSNDYSFIRYTRTYLQFQFOALCOAKHGEPLHKCDISNSTEAGOKLFNML 540
DB 499 DPASLFHVSNDYSFIRYTRTYLQFQFOALCOAKHGEPLHKCDISNSTEAGOKLFNML 558
QY 541 RLKGRSEPTLALENVGAKNNVRLNTEPELFTWLKDONKNSFWGSTDMSPYADQSI 600
DB 559 RLKGRSEPTLALENVGAKNNVRLNTEPELFTWLKDONKNSFWGSTDMSPYADQSI 618
QY 601 KVRISLSALGDKAYEMNDNEMTLFRSSVAYAMQYFLKYNOMILFGEEDVAVANLKR 660
DB 619 KVRISLSALGDKAYEMNDNEMTLFRSSVAYAMQYFLKYNOMILFGEEDVAVANLKR 678
QY 661 ISFNEFVAPKNSDIIIPREVEKAIKMSRSRINDAFLRNDNSLEFLGIQPTLGPNOQP 720
DB 679 ISFNEFVAPKNSDIIIPREVEKAIKMSRSRINDAFLRNDNSLEFLGIQPTLGPNOQP 738

RESULT 2

US-10-158-847-142
; Sequence 142, Application US/10158847
; Publication No. US2003009157A1
; GENERAL INFORMATION:
; APPLICANT: Tom Parry et al.
; TITLE OF INVENTION: Method and Compositions for Modulating ACE-2 Activity
; FILE REFERENCE: P557
; CURRENT APPLICATION NUMBER: US/10/158,847
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 60/295,004
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 158
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 142
; LENGTH: 805
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-158-847-142

Query Match 100.0%; Score 3869; DB 9; Length 805;

Best Local Similarity 100.0%; Pred. No. 1,9e-312;

Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIEQAKTFLDKNEHEADLFYQSSLASWNTNTITEENVQNNAGKWSAFLKEOST 60
DB 19 STIEQAKTFLDKNEHEADLFYQSSLASWNTNTITEENVQNNAGKWSAFLKEOST 78
QY 61 LAQWYPIQOEIONTVAKIQALAOQNGSSVSEDSKSKRLNTLMTSTYSTGVCPNDP 120
DB 79 LAQWYPIQOEIONTVAKIQALAOQNGSSVSEDSKSKRLNTLMTSTYSTGVCPNDP 138
QY 121 QECLEBGLNEIANSIDYERLMAWESWSEVQKOLRLPYEYVVLAKNEMARANYED 180
DB 139 QECLEBGLNEIANSIDYERLMAWESWSEVQKOLRLPYEYVVLAKNEMARANYED 198

QY 181 YGDYWRGDEYVNGVGYDYSRGQLIEDYEHTEFEIKPLYEHLHAYVRAKLMAAYPSYISP 240
DB 199 YGDYWRGDEYVNGVGYDYSRGQLIEDYEHTEFEIKPLYEHLHAYVRAKLMAAYPSYISP 258
QY 241 IGCPLAHLGDMGGRFTNLYSLTPFGOKPNDIVTDAMVDQAMDQORIFKEAEKFFVS 300
DB 259 IGCPLAHLGDMGGRFTNLYSLTPFGOKPNDIVTDAMVDQAMDQORIFKEAEKFFVS 318
QY 301 GLPNTGFWENSMLTDPGNVOKAVCHPTAMDLCGDRILMCTRYTMDDFLTAHHEMGH 360
DB 319 GLPNTGFWENSMLTDPGNVOKAVCHPTAMDLCGDRILMCTRYTMDDFLTAHHEMGH 378
QY 361 IQYDMAYAAOPFLLRNGANGFHEAVGEIMSLAATPKHLKISGLSPDFQEDNETEINF 420
DB 379 IQYDMAYAAOPFLLRNGANGFHEAVGEIMSLAATPKHLKISGLSPDFQEDNETEINF 438
QY 421 LKQALITVGLPTFTYMLEKRWNVFEGEIPKQDMKKWEMKREITGVVEVPDDETYC 480
DB 439 LKQALITVGLPTFTYMLEKRWNVFEGEIPKQDMKKWEMKREITGVVEVPDDETYC 498
QY 481 DPASLFHVSNDYSFIRYTRTYLQFQFOALCOAKHGEPLHKCDISNSTEAGOKLFNML 540
DB 499 DPASLFHVSNDYSFIRYTRTYLQFQFOALCOAKHGEPLHKCDISNSTEAGOKLFNML 558
QY 541 RLKGRSEPTLALENVGAKNNVRLNTEPELFTWLKDONKNSFWGSTDMSPYADQSI 600
DB 559 RLKGRSEPTLALENVGAKNNVRLNTEPELFTWLKDONKNSFWGSTDMSPYADQSI 618
QY 601 KVRISLSALGDKAYEMNDNEMTLFRSSVAYAMQYFLKYNOMILFGEEDVAVANLKR 660
DB 619 KVRISLSALGDKAYEMNDNEMTLFRSSVAYAMQYFLKYNOMILFGEEDVAVANLKR 678
QY 661 ISFNEFVAPKNSDIIIPREVEKAIKMSRSRINDAFLRNDNSLEFLGIQPTLGPNOQP 720
DB 679 ISFNEFVAPKNSDIIIPREVEKAIKMSRSRINDAFLRNDNSLEFLGIQPTLGPNOQP 738

RESULT 3

US-10-114-893-86
; Sequence 86, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalley, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; EARLIER APPLICATION NUMBER: 09/413,232
; EARLIER FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 86
; LENGTH: 805
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-893-86

Query Match 99.9%; Score 3865; DB 9; Length 805;

Best Local Similarity 99.9%; Pred. No. 4.1e-312;

Matches 719; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 STIEQAKTFLDKRNEHAEELFYOSSLASMYNTNITEENVONNANNGDKMSAFLEKOST 60
Db 19 STIEQAKTFLDKRNEHAEELFYOSSLASMYNTNITEENVONNANNGDKMSAFLEKOST 78
OY 61 LAQWYPLQEIOMLVKLOLQALQONGSSVLSSEKSKRLNTLNTMTSTTSTGKVCNPNP 120
Db 79 LAQWYPLQEIOMLVKLOLQALQONGSSVLSSEKSKRLNTLNTMTSTTSTGKVCNPNP 138
OY 121 QECILLEPGLNEIMANSIDYNERLMAWESMRSEVQKQRLPYEERYVLKKNEMARANYED 180
Db 139 QECILLEPGLNEIMANSIDYNERLMAWESMRSEVQKQRLPYEERYVLKKNEMARANYED 198
OY 181 YGDWRGDIYEVNGVDGYSRGQLEIEVEHTFEETIKLYEHLAAYRAKLMAVPSYISP 240
Db 199 YGDWRGDIYEVNGVDGYSRGQLEIEVEHTFEETIKLYEHLAAYRAKLMAVPSYISP 258
OY 241 IGCPLPAHLGDMGCFWTNLVSLVPPGQKPNIDVTAMVDOAMDORIKKEKEKFFVSV 300
Db 259 IGCPLPAHLGDMGCFWTNLVSLVPPGQKPNIDVTAMVDOAMDORIKKEKEKFFVSV 318
OY 301 GLPMTQGFENSMKLDPGNVQKAVCHPTAMDLSKGDPRILMCTKYVMDPFLTAHHEMGH 360
Db 319 GLPMTQGFENSMKLDPGNVQKAVCHPTAMDLSKGDPRILMCTKYVMDPFLTAHHEMGH 378
OY 361 IYQDMAVYAAQPFLLRNGANGEGHEAVGEIMSLSAATPKHLKSLGILSPDEQDNETEINF 420
Db 379 IYQDMAVYAAQPFLLRNGANGEGHEAVGEIMSLSAATPKHLKSLGILSPDEQDNETEINF 438
OY 421 LKQALITVGLTPTVTLERKRMWVFGELIPKDOMKKWMEKRELYGVVEPHEDETC 480
Db 439 LKQALITVGLTPTVTLERKRMWVFGELIPKDOMKKWMEKRELYGVVEPHEDETC 498
OY 481 DPASLHVSNDSYFIRYTRTYLQFOFOALCOAAHCEGPLKCDISNSTEAGOKLFNML 540
Db 499 DPASLHVSNDSYFIRYTRTYLQFOFOALCOAAHCEGPLKCDISNSTEAGOKLFNML 558
OY 541 RLKGSBPWTLEENYVGAKNMVRPLNTEPELFTLKDQNNKSPFGSTDMSPYADOSI 600
Db 559 RLKGSBPWTLEENYVGAKNMVRPLNTEPELFTLKDQNNKSPFGSTDMSPYADOSI 618
OY 601 KVAISIKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVKKNOMILFGEEDVRYANLKPR 660
Db 619 KVAISIKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVKKNOMILFGEEDVRYANLKPR 678
OY 661 ISFNFVTAPKKNVSDIIPRTVEKKAIRMSRSRINDAFLRNDLSLEFLGIQPTLGPNOQP 720
Db 679 ISFNFVTAPKKNVSDIIPRTVEKKAIRMSRSRINDAFLRNDLSLEFLGIQPTLGPNOQP 738

```

RESULT 4
US-09-969-384-25
Sequence 25, Application US/09969384
Publication No. US20020192749A1

GENERAL INFORMATION:

APPLICANT: Moore, et al.
TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
FILE REFERENCE: PTO55P1
CURRENT APPLICATION NUMBER: US/09/969,384
CURRENT FILING DATE: 2001-10-03
PRIOR APPLICATION NUMBER: PCT/US01/10542
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 60/236,384
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/194,118
PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 25
LENGTH: 681

TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE

LOCATION: (219)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (240)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (499)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-969-384-25

Query Match 93.6%; Score 3620; DB 9; Length 681;
Best Local Similarity 99.3%; Pred. No. 7e-292;
Matches 672; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

OY 44 MNNAGDKMSAFLEKOSTLAQMTPLQEIOMLVKLOLQALQONGSSVLSSEKSKRLNTLNT 103
Db 1 MNNAGDKMSAFLEKOSTLAQMTPLQEIOMLVKLOLQALQONGSSVLSSEKSKRLNTLNT 60
OY 104 TMTSTYTGKVCNPNPQECILLEPGLNEIMANSIDYNERLMAWESMRSEVQKQRLPYE 163
Db 61 TMTSTYTGKVCNPNPQECILLEPGLNEIMANSIDYNERLMAWESMRSEVQKQRLPYE 120
OY 164 EYVVLKKNEMARANYEDYGDWRGDIYEVNGVDGYSRGQLEIEVEHTFEETIKLYEHL 223
Db 121 EYVVLKKNEMARANYEDYGDWRGDIYEVNGVDGYSRGQLEIEVEHTFEETIKLYEHL 180
OY 224 AYVRKAKMAVPSYISPIGCLPFAHLGDMGCFWTNLVSLVPPGQKPNIDVTAMVDOA 283
Db 181 AYVRKAKMAVPSYISPIGCLPFAHLGDMGCFWTNLVSLVPPGQKPNIDVTAMVDOA 240
OY 284 WDAQRIFEAEKEFFSVGLPNMTQGFENSMKLDPGNVQKAVCHPTAMDLSKGDPRILMCT 343
Db 241 WDAQRIFEAEKEFFSVGLPNMTQGFENSMKLDPGNVQKAVCHPTAMDLSKGDPRILMCT 300
OY 344 TKVTMDPFLTAHHEMGHIOYDMAVYAAQPFLLRNGANGEGHEAVGEIMSLSAATPKHLKSI 403
Db 301 TKVTMDPFLTAHHEMGHIOYDMAVYAAQPFLLRNGANGEGHEAVGEIMSLSAATPKHLKSI 360
OY 404 GLSPDEQDNETEINFLLKQALITVGLTPTVTLERKRMWVFGELIPKDOMKKWMEKRE 463
Db 361 GLSPDEQDNETEINFLLKQALITVGLTPTVTLERKRMWVFGELIPKDOMKKWMEKRE 420
OY 464 REIYGVEPVPHEDETCDPASLHVSNDSYFIRYTRTYLQFOFOALCOAAHCEGPLK 523
Db 421 REIYGVEPVPHEDETCDPASLHVSNDSYFIRYTRTYLQFOFOALCOAAHCEGPLK 480
OY 524 CDISNSTEAGOKLFNMLRLKGSBPWTLEENYVGAKNMVRPLNTEPELFTLKDQNNK 583
Db 481 CDISNSTEAGOKLFNMLRLKGSBPWTLEENYVGAKNMVRPLNTEPELFTLKDQNNK 540
OY 584 SFVGMSTDMSPYADOSIKVAISIKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVK 643
Db 541 SFVGMSTDMSPYADOSIKVAISIKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVK 600
OY 644 MILEGEEDVRYANLKPRISFNFVTAPKKNVSDIIPRTVEKKAIRMSRSRINDAFLRND 703
Db 601 MILEGEEDVRYANLKPRISFNFVTAPKKNVSDIIPRTVEKKAIRMSRSRINDAFLRND 660
OY 704 LEFLGIQPTLGPNOQP 720
Db 661 LEFLGIQPTLGPNOQP 677

```

RESULT 5
US-10-158-847-140
Sequence 140, Application US/10158847
Publication No. US20030091557A1
GENERAL INFORMATION:
APPLICANT: Tom Parry et al.
TITLE OF INVENTION: Method and Compositions for Modulating ACE-2 Activity
FILE REFERENCE: PFS57
CURRENT APPLICATION NUMBER: US/10/158,847
CURRENT FILING DATE: 2002-06-03

PRIOR APPLICATION NUMBER: 60/295,004
 PRIOR FILING DATE: 2001-06-04
 NUMBER OF SEQ ID NOS: 158
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 140
 LENGTH: 681
 TYPE: PRT
 ORGANISM: homo sapiens
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (219)..(219)
 OTHER INFORMATION: xaa equals any amino acid
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (240)..(240)
 OTHER INFORMATION: xaa equals any amino acid
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (499)..(499)
 OTHER INFORMATION: xaa equals any amino acid
 US-10-158-847-140

Query Match 93.6%; Score 3620; DB 9; Length 681;
 Best Local Similarity 99.3%; Pred. No. 7e-292;

Matches 672; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

44 MNNAGDKMSAFLEKQSTLAQMYPLQEIONTLVKIQLOALQONGSSVLSSEKSKRLNTLN 103
 1 MNNAGDKMSAFLEKQSTLAQMYPLQEIONTLVKIQLOALQONGSSVLSSEKSKRLNTLN 60
 104 TMSITSTGKVCNPNPOECILLPGLNEIMANSIDYNERLWAMESRSSEVGKQRLPYE 163
 61 TMSITSTGKVCNPNPOECILLPGLNEIMANSIDYNERLWAMESRSSEVGKQRLPYE 120
 164 EYVLKNEMARANHYEDYDWRGDEYVNGVDYDYSRGQLIEDVEHTFEIKPLYEHLH 223
 121 EYVLKNEMARANHYEDYDWRGDEYVNGVDYDYSRGQLIEDVEHTFEIKPLYEHLH 180
 224 AYVRKLMANAYPSYISPIGCLPAHLGDMGRFTNLXSLTVPGOKPNDIVTDAMVDOA 283
 181 AYVRKLMANAYPSYISPIGCLPAHLGDMGRFTNLXSLTVPGOKPNDIVTDAMVDOX 240
 284 WDAORIFKEAEKEFEFVSGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDLGDFRILMC 343
 241 WDAORIFKEAEKEFEFVSGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDLGDFRILMC 300
 344 TKTMDLFLTAHHEMHIQYDMAVAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSI 403
 301 TKTMDLFLTAHHEMHIQYDMAVAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSI 360
 404 GLSPDFQEDNTEINFLKQALITVGLPFTYMLEKRMWVFKEIPIKDDMKKWEK 463
 361 GLSPDFQEDNTEINFLKQALITVGLPFTYMLEKRMWVFKEIPIKDDMKKWEK 420
 464 RELVGVVEVPYPHDETCDPASLFHVSNDYSFIRYTRTLXQFOFQALQAAKHGEPILK 523
 421 RELVGVVEVPYPHDETCDPASLFHVSNDYSFIRYTRTLXQFOFQALQAAKHGEPILK 480
 524 CDISNSTEAGQKLFNNMLRLGKSEPTLALENVVGAKNMVRPLNTEFPLTWLKDQKN 583
 481 CDISNSTEAGQKLFNNMLRLGKSEPTLALENVVGAKNMVRPLNTEFPLTWLKDQKN 540
 584 SPVGMSTWSPYADQSIKVRISLSKALGDKAYTENDNENKTLFRSSVAYAMROFLKVKNO 643
 541 SPVGMSTWSPYADQSIKVRISLSKALGDKAYTENDNENKTLFRSSVAYAMROFLKVKNO 600
 644 MILEGEEDVYANAKPRISNFETAPKNSDIIIPREYKAIKMSRSRINDAFRLNDS 703
 601 MILEGEEDVYANAKPRISNFETAPKNSDIIIPREYKAIKMSRSRINDAFRLNDS 660
 704 LEFLGIQPTLGPNNOP 720
 661 LEFLGIQPTLGPNNOP 677

RESULT 6
 US-09-969-384-13
 Sequence 13, Application US/09969384
 Publication No. US20020192749A1
 GENERAL INFORMATION:
 APPLICANT: Moore, et al.
 TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
 FILE REFERENCE: P055PI
 CURRENT APPLICATION NUMBER: US/09/969,384
 PRIOR FILING DATE: 2001-10-03
 PRIOR APPLICATION NUMBER: PCT/US01/10542
 PRIOR FILING DATE: 2001-04-02
 PRIOR APPLICATION NUMBER: 60/236,384
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/194,118
 PRIOR FILING DATE: 2000-04-03
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 13
 LENGTH: 711
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (219)
 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (499)
 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
 US-09-969-384-13

Query Match 93.6%; Score 3620; DB 9; Length 711;
 Best Local Similarity 99.3%; Pred. No. 7.5e-292;

Matches 672; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

44 MNNAGDKMSAFLEKQSTLAQMYPLQEIONTLVKIQLOALQONGSSVLSSEKSKRLNTLN 103
 1 MNNAGDKMSAFLEKQSTLAQMYPLQEIONTLVKIQLOALQONGSSVLSSEKSKRLNTLN 60
 104 TMSITSTGKVCNPNPOECILLPGLNEIMANSIDYNERLWAMESRSSEVGKQRLPYE 163
 61 TMSITSTGKVCNPNPOECILLPGLNEIMANSIDYNERLWAMESRSSEVGKQRLPYE 120
 164 EYVLKNEMARANHYEDYDWRGDEYVNGVDYDYSRGQLIEDVEHTFEIKPLYEHLH 223
 121 EYVLKNEMARANHYEDYDWRGDEYVNGVDYDYSRGQLIEDVEHTFEIKPLYEHLH 180
 224 AYVRKLMANAYPSYISPIGCLPAHLGDMGRFTNLXSLTVPGOKPNDIVTDAMVDOA 283
 181 AYVRKLMANAYPSYISPIGCLPAHLGDMGRFTNLXSLTVPGOKPNDIVTDAMVDOX 240
 284 WDAORIFKEAEKEFEFVSGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDLGDFRILMC 343
 241 WDAORIFKEAEKEFEFVSGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDLGDFRILMC 300
 344 TKTMDLFLTAHHEMHIQYDMAVAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSI 403
 301 TKTMDLFLTAHHEMHIQYDMAVAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSI 360
 404 GLSPDFQEDNTEINFLKQALITVGLPFTYMLEKRMWVFKEIPIKDDMKKWEK 463
 361 GLSPDFQEDNTEINFLKQALITVGLPFTYMLEKRMWVFKEIPIKDDMKKWEK 420
 464 RELVGVVEVPYPHDETCDPASLFHVSNDYSFIRYTRTLXQFOFQALQAAKHGEPILK 523
 421 RELVGVVEVPYPHDETCDPASLFHVSNDYSFIRYTRTLXQFOFQALQAAKHGEPILK 480
 524 CDISNSTEAGQKLFNNMLRLGKSEPTLALENVVGAKNMVRPLNTEFPLTWLKDQKN 583

Db 481 CDISNSTEAGOKLFNMLRGKSEPTWLTALENVGAKNNVPLNTYFETLTKDQKN 540
QY 584 SEFGNSTDMSPYADOSIKVRSLSKALDRAKAYENDNEMTLFRSSVAYAMROYFLKVKNO 643
Db 541 SEFGNSTDMSPYADOSIKVRSLSKALDRAKAYENDNEMTLFRSSVAYAMROYFLKVKNO 600
QY 644 MILGGEEDYRVANLKPRISENFVYAPKNVSDIIPREVEKATIMRSRINDAFRLDMS 703
Db 601 MILGGEEDYRVANLKPRISENFVYAPKNVSDIIPREVEKATIMRSRINDAFRLDMS 660
QY 704 LEFGIOPTLGPPNPP 720
Db 661 LEFGIOPTLGPPNPP 677

RESULT 7
US-10-158-847-138
; Sequence 138, Application US/10158847
; Publication No. US20030091557A1
; GENERAL INFORMATION:
; APPLICANT: Tom Parry et al.
; TITLE OF INVENTION: Method and Compositions for Modulating ACE-2 Activity
; FILE REFERENCE: PE557
; CURRENT APPLICATION NUMBER: US/10/158,847
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 60/295,004
; NUMBER OF SEQ ID NOS: 158
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 138
; LENGTH: 711
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (219)..(219)
; OTHER INFORMATION: xaa equals any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (240)..(240)
; OTHER INFORMATION: xaa equals any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (499)..(499)
; OTHER INFORMATION: xaa equals any amino acid
US-10-158-847-138

Query Match 93.6%; Score 3620; DB 9; Length 711;
Best Local Similarity 99.3%; Pred. No. 7, 5e-292;
Matches 672; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 44 MNAGDKMSAFLEKQSTLAOMYPLQEIOMLTNYKLOLQALQOQSSVLSSEKSKRLNTLN 103
Db 1 MNAGDKMSAFLEKQSTLAOMYPLQEIOMLTNYKLOLQALQOQSSVLSSEKSKRLNTLN 60
QY 104 TMSITISGKVCNPNPOECILLLEPGLNEIMANSIDYNERLMAWESMSSEVGKOLRPLYE 163
Db 61 TMSITISGKVCNPNPOECILLLEPGLNEIMANSIDYNERLMAWESMSSEVGKOLRPLYE 120
QY 164 EYVVLKNEKARAHNEEDYGDYWRGDEYVNGVDGIDYSGQLEIEVEHTEFEIKPLYEHL 223
Db 121 EYVVLKNEKARAHNEEDYGDYWRGDEYVNGVDGIDYSGQLEIEVEHTEFEIKPLYEHL 180
QY 224 AYVRAKLMAVPSYISPIGCLPAHLGDMGRFTNTLSLVPPGOKPNIDVDYDAMVDA 263
Db 181 AYVRAKLMAVPSYISPIGCLPAHLGDMGRFTNTLSLVPPGOKPNIDVDYDAMVDA 240
QY 284 WDAQIFKEAEKFPFVSGLPMWOGFENSLTDPGAVOAVCHPTAMDLGKGDFTILMC 343
Db 241 WDAQIFKEAEKFPFVSGLPMWOGFENSLTDPGAVOAVCHPTAMDLGKGDFTILMC 300
QY 344 TKVTMDFLTAHHEMGIQYDMAVAAQFLLRNGANGCFEAVGEIMSLAATPKHLKSI 403

Db 301 TKVTMDFLTAHHEMGIQYDMAVAAQFLLRNGANGCFEAVGEIMSLAATPKHLKSI 360
QY 404 GILSDPQONDNTEINFLKQALTVGLPTTYMLEKRWVVFGEIIPKQOMKRWEMK 463
Db 361 GILSDPQONDNTEINFLKQALTVGLPTTYMLEKRWVVFGEIIPKQOMKRWEMK 420
QY 464 REIVGVSEVPNDETCYCPASLFHSNDYSIRYRTYTRTLOFQOQALCOAAHHEGLK 523
Db 421 REIVGVSEVPNDETCYCPASLFHSNDYSIRYRTYTRTLOFQOQALCOAAHHEGLK 480
QY 524 CDISNSTEAGOKLFNMLRGKSEPTWLTALENVGAKNNVPLNTYFETLTKDQKN 583
Db 481 CDISNSTEAGOKLFNMLRGKSEPTWLTALENVGAKNNVPLNTYFETLTKDQKN 540
QY 584 SEFGNSTDMSPYADOSIKVRSLSKALDRAKAYENDNEMTLFRSSVAYAMROYFLKVKNO 643
Db 541 SEFGNSTDMSPYADOSIKVRSLSKALDRAKAYENDNEMTLFRSSVAYAMROYFLKVKNO 600
QY 644 MILGGEEDYRVANLKPRISENFVYAPKNVSDIIPREVEKATIMRSRINDAFRLDMS 703
Db 601 MILGGEEDYRVANLKPRISENFVYAPKNVSDIIPREVEKATIMRSRINDAFRLDMS 660
QY 704 LEFGIOPTLGPPNPP 720
Db 661 LEFGIOPTLGPPNPP 677

RESULT 8
US-09-978-385-6
; Sequence 6, Application US/09978385
; Patent No. US2002017211A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Petrie, Charles
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ZAC2: A HUMAN METALLOENZYME
; FILE REFERENCE: 99-24C1
; CURRENT APPLICATION NUMBER: US/09/978,385
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/133,952
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 60/151,181
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/563,516
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 805
; TYPE: PRT
; ORGANISM: Mouse
US-09-978-385-6

Query Match 84.0%; Score 3251; DB 9; Length 805;
Best Local Similarity 83.1%; Pred. No. 3, 8e-261;
Matches 598; Conservative 49; Mismatches 73; Indels 0; Gaps 0;

QY 1 STIEBQAKTELDKEHEADLFYQSSLASMNTNITEENVQNNMNGDKMSAFLEKQST 60
Db 19 STIEBQAKTELDKEHEADLFYQSSLASMNTNITEENVQNNMNGDKMSAFLEKQST 78
QY 61 LAOMYPLQEIOMLTNYKLOLQALQOQSSVLSSEKSKRLNTLNMTISYTGVCNPN 120
Db 79 TAQSFSLQEIOPPIIKRQLOALQOQSSVLSADKNKQALNTLNMTISYTGVCNPN 138
QY 121 QCCILLEGNEIMANSIDYNERLMAWESMSSEVGKOLRPLYEYVVLKNEKARAHNEED 180
Db 139 QCCILLEGNEIMANSIDYNERLMAWESMSSEVGKOLRPLYEYVVLKNEKARAHNEED 198
QY 181 YGDYWRGDEYVNGVDGIDYSRQGLEIEVEHTEFEIKPLYEHLTAAYPAKMAVPSYIS 240

Db 199 YGDIWRGDYEAEGADGYNRNQLLIEDVETFAEIKPLYEHLHAYVRKLMJTYSTSP 258
QY 241 IGCIPAHLLDGMGCRFTNLXSLTVPFGOKPNIDVTDAMVDAQMRIFKEAEKEFVS 300
QY 259 TGCIPAHLLDGMGCRFTNLXSLTVPFAOKPNIDVTDAMNOCWDAERIFQAEKEFVS 318
QY 301 GLPNMTOGFEMNSMLTDPGNVQAVCHPTAMDLGKDEFRILMCTKYTMDFLTAHHEGH 360
Db 319 GLPHMTOGFEMNSMLTDPADGRKYVCHPTAMDIGHDFRILMCTKYTMDFLTAHHEGH 378
QY 361 IOYMAVRAOPEFLLRNANGFHEAVGEIMSLAATPKHLKSGLLSPFOEDNETEINF 420
Db 379 IOYMAVRAOPEFLLRNANGFHEAVGEIMSLAATPKHLKSGLLSPFOEDNETEINF 438
QY 421 LKQALITVGLPTPTMLEKRWAVFKEGELPKDQMKKWMEMKREIVGVEPLPHDETC 480
Db 439 LKQALITVGLPTPTMLEKRWAVFKEGELPKDQMKKWMEMKREIVGVEPLPHDETC 498
QY 481 DPASLEFHSNDYSFIRYTRITLYOFQOALCOAKHGEPLHKCDISNSTEAGOKLFENML 540
Db 499 DPASLEFHSNDYSFIRYTRITLYOFQOALCOAKHGEPLHKCDISNSTEAGOKLFENML 558
QY 541 RIGKSEPTLLENVYGAKNANVRPLNTEPELFTWLDKONKSEFGWSTDSWSPYADQSI 600
Db 559 SIGNEPPTLLENVYGAKNANVRPLNTEPELFTWLDKONKSEFGWSTDSWSPYADQSI 618
QY 601 KVRISLKSALGDKAYEMNDENMYLFRSSVAYAMQYFLKYNQMLLGEEDVAVANLKR 660
Db 619 KVRISLKSALGANAEWNNEMFLFRSSVAYAMKRYSLIKNOVPLEEDVAVSOLKPR 678
QY 661 ISFNFVYAPKRVSDIIPRETEVEKIRMSRINDAFRLNDSLEFLGIOTPLGPPNQP 720
Db 679 VSFFYFVSPQVSDVIRSEVEDAIRSGRINDVGLNDSLEFLGIHTLEPPYQPP 738

RESULT 9

US-09-978-385-9
Sequence 9, Application US/09978385
Patent No. US2002017211A1
GENERAL INFORMATION:
APPLICANT: Piddington, Christopher S.
APPLICANT: Petrie, Charles
APPLICANT: Shoemaker, Kimberly E.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME
FILE REFERENCE: 99-24C1
CURRENT APPLICATION NUMBER: US/09/978,385
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 60/133,952
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: 60/151,181
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: 09/563,516
PRIOR FILING DATE: 2000-05-03
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 9
LENGTH: 805
TYPE: PR
ORGANISM: Mouse
US-09-978-385-9

Query Match 83.6%; Score 3233; DB 9; Length 805;
Best Local Similarity 82.8%; Pred. No. 1.2e-259;
Matches 596; Conservative 49; Mismatches 75; Indels 0; Gaps 0;

QY 1 STIEQAKTFLDKNNHAEEDLFYOSSLASMYNTNTTEENVOMNNAAGDKMSAFLEKOST 60
Db 19 SLTEENAKTFLNNNOEAEDLSSTOSSLASMYNTNTTEENAOQMSAAKMSAFLEKOST 78
QY 61 LAQWYPLQEIQNTLVKQLQALQOQNGSSVLSSEKSKRLNTLITMSTIYSTGKVCNDP 120
Db 79 TQSFSTQEIOTPIIKKQALQOQSSSALSADKNKQMLNTLITMSTIYSTGKVCNDP 138

QY 121 OECILLEPGNEITANSIDYNEERLNAWESRSEVGKOLRPLYEEVVLKNEMARANHYED 180
Db 139 OECILLEPGNEIDEMASTIDYNSRLNMEGRRAEVGOLRPLYEEVVLKNEMARANHYED 198
QY 181 YGDIWRGDYEAEGADGYNRNQLLIEDVETFAEIKPLYEHLHAYVRKLMJTYSTSP 240
Db 199 YGDIWRGDYEAEGADGYNRNQLLIEDVETFAEIKPLYEHLHAYVRKLMJTYSTSP 258
QY 241 IGCIPAHLLDGMGCRFTNLXSLTVPFGOKPNIDVTDAMVDAQMRIFKEAEKEFVS 300
Db 259 TGCIPAHLLDGMGCRFTNLXSLTVPFAOKPNIDVTDAMNOCWDAERIFQAEKEFVS 318
QY 301 GLPNMTOGFEMNSMLTDPGNVQAVCHPTAMDLGKDEFRILMCTKYTMDFLTAHHEGH 360
Db 319 GLPHMTOGFEMNSMLTDPADGRKYVCHPTAMDIGHDFRILMCTKYTMDFLTAHHEGH 378
QY 361 IOYMAVRAOPEFLLRNANGFHEAVGEIMSLAATPKHLKSGLLSPFOEDNETEINF 420
Db 379 IOYMAVRAOPEFLLRNANGFHEAVGEIMSLAATPKHLKSGLLSPFOEDNETEINF 438
QY 421 LKQALITVGLPTPTMLEKRWAVFKEGELPKDQMKKWMEMKREIVGVEPLPHDETC 480
Db 439 LKQALITVGLPTPTMLEKRWAVFKEGELPKDQMKKWMEMKREIVGVEPLPHDETC 498
QY 481 DPASLEFHSNDYSFIRYTRITLYOFQOALCOAKHGEPLHKCDISNSTEAGOKLFENML 540
Db 499 DPASLEFHSNDYSFIRYTRITLYOFQOALCOAKHGEPLHKCDISNSTEAGOKLFENML 558
QY 541 RIGKSEPTLLENVYGAKNANVRPLNTEPELFTWLDKONKSEFGWSTDSWSPYADQSI 600
Db 559 SIGNEPPTLLENVYGAKNANVRPLNTEPELFTWLDKONKSEFGWSTDSWSPYADQSI 618
QY 601 KVRISLKSALGDKAYEMNDENMYLFRSSVAYAMQYFLKYNQMLLGEEDVAVANLKR 660
Db 619 KVRISLKSALGANAEWNNEMFLFRSSVAYAMKRYSLIKNOVPLEEDVAVSOLKPR 678
QY 661 ISFNFVYAPKRVSDIIPRETEVEKIRMSRINDAFRLNDSLEFLGIOTPLGPPNQP 720
Db 679 VSFFYFVSPQVSDVIRSEVEDAIRSGRINDVGLNDSLEFLGIHTLEPPYQPP 738

RESULT 10

US-10-028-072-72
Sequence 72, Application US/10028072
Publication No. US20030004311A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Pilyaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltisen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OF INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028,072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113

1	PRIOR APPLICATION NUMBER: 60/0696934
2	PRIOR FILING DATE: 1997-12-16
3	PRIOR APPLICATION NUMBER: 60/072320
4	PRIOR FILING DATE: 1998-01-23
5	PRIOR APPLICATION NUMBER: 60/073612
6	PRIOR FILING DATE: 1998-02-04
7	PRIOR APPLICATION NUMBER: 60/074086
8	PRIOR FILING DATE: 1998-02-09
9	PRIOR APPLICATION NUMBER: 60/074092
10	PRIOR FILING DATE: 1998-02-09
11	PRIOR APPLICATION NUMBER: 60/077791
12	PRIOR FILING DATE: 1998-03-12
13	PRIOR APPLICATION NUMBER: 60/078910
14	PRIOR FILING DATE: 1998-03-20
15	PRIOR APPLICATION NUMBER: 60/079294
16	PRIOR FILING DATE: 1998-03-25
17	PRIOR APPLICATION NUMBER: 60/079663
18	PRIOR FILING DATE: 1998-02-27
19	PRIOR APPLICATION NUMBER: 60/079728
20	PRIOR FILING DATE: 1998-03-27
21	PRIOR APPLICATION NUMBER: 60/080165
22	PRIOR FILING DATE: 1998-03-31
23	PRIOR APPLICATION NUMBER: 60/081203
24	PRIOR FILING DATE: 1998-04-09
25	PRIOR APPLICATION NUMBER: 60/081229
26	PRIOR FILING DATE: 1998-04-09
27	PRIOR APPLICATION NUMBER: 60/081695
28	PRIOR FILING DATE: 1998-04-14
29	PRIOR APPLICATION NUMBER: 60/081817
30	PRIOR FILING DATE: 1998-04-15
31	PRIOR APPLICATION NUMBER: 60/081818
32	PRIOR FILING DATE: 1998-04-15
33	PRIOR APPLICATION NUMBER: 60/082999
34	PRIOR FILING DATE: 1998-04-24
35	PRIOR APPLICATION NUMBER: 60/083322
36	PRIOR FILING DATE: 1998-04-28
37	PRIOR APPLICATION NUMBER: 60/083545
38	PRIOR FILING DATE: 1998-04-29
39	PRIOR APPLICATION NUMBER: 60/084600
40	PRIOR FILING DATE: 1998-05-07
41	PRIOR APPLICATION NUMBER: 60/084627
42	PRIOR FILING DATE: 1998-05-07
43	PRIOR APPLICATION NUMBER: 60/084637
44	PRIOR FILING DATE: 1998-05-07
45	PRIOR APPLICATION NUMBER: 60/085149
46	PRIOR FILING DATE: 1998-05-12
47	PRIOR APPLICATION NUMBER: 60/085323
48	PRIOR FILING DATE: 1998-05-13
49	PRIOR APPLICATION NUMBER: 60/085338
50	PRIOR FILING DATE: 1998-05-13
51	PRIOR APPLICATION NUMBER: 60/085339
52	PRIOR FILING DATE: 1998-05-13
53	PRIOR APPLICATION NUMBER: 60/085579
54	PRIOR FILING DATE: 1998-05-15
55	PRIOR APPLICATION NUMBER: 60/085697
56	PRIOR FILING DATE: 1998-05-15
57	PRIOR APPLICATION NUMBER: 60/085704
58	PRIOR FILING DATE: 1998-05-15
59	PRIOR APPLICATION NUMBER: 60/086414
60	PRIOR FILING DATE: 1998-05-22
61	PRIOR APPLICATION NUMBER: 60/086430
62	PRIOR FILING DATE: 1998-05-22
63	PRIOR APPLICATION NUMBER: 60/087106
64	PRIOR FILING DATE: 1998-05-28
65	PRIOR APPLICATION NUMBER: 60/088026
66	PRIOR FILING DATE: 1998-06-04
67	PRIOR APPLICATION NUMBER: 60/088730
68	PRIOR FILING DATE: 1998-06-10
69	PRIOR APPLICATION NUMBER: 60/088741
70	PRIOR FILING DATE: 1998-06-10
71	PRIOR APPLICATION NUMBER: 60/088810
72	PRIOR FILING DATE: 1998-06-10
73	PRIOR APPLICATION NUMBER: 60/088858

PRIOR FILING DATE: 19/98-06-11
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089559
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090538
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090863
 PRIOR FILING DATE: 1998-06-28
 PRIOR APPLICATION NUMBER: 60/091360
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091519
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091982
 PRIOR FILING DATE: 1998-07-07

Query Match 74.9%; Score 2897; DB 9; Length 555;
 Best Local Similarity 99.8%; Pred. No. 5.3e-232;
 Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 STEEOATFLDKFHEAEDLFYOSLASMYNTNITEENVONNNAAGKMSAFLEKOST 60
 19 STEEOATFLDKFHEAEDLFYOSLASMYNTNITEENVONNNAAGKMSAFLEKOST 78
 61 LAQMTPLDQIOMLVKLOLQALQONGSSVLESDSKRLNTLNTMSTIYSTGVCNPNP 120
 79 LAQMTPLDQIOMLVKLOLQALQONGSSVLESDSKRLNTLNTMSTIYSTGVCNPNP 138
 121 QECILLEPGLNEIMANSIDYNERLMAWESRSEVSKOLRPLYEEVVLKEMARANYED 180
 139 QECILLEPGLNEIMANSIDYNERLMAWESRSEVSKOLRPLYEEVVLKEMARANYED 198
 181 YGDYWRGDEYVNGVGYDYSRGQLIEDVEHFEETIKPLYEHLHAYRAKLMNAPSYISP 240
 199 YGDYWRGDEYVNGVGYDYSRGQLIEDVEHFEETIKPLYEHLHAYRAKLMNAPSYISP 258
 241 IGLPRAHLIGDMWGRFWTNLSLVFPGOKPNIDVTDAVDAQADRIEFAEKFFVS 300
 259 IGLPRAHLIGDMWGRFWTNLSLVFPGOKPNIDVTDAVDAQADRIEFAEKFFVS 318
 301 GLPNMTOGFWENSMILDPGNVQKAVCHPTAMDGLGKDFRIIMCTKVTMDDELTAAHHEGH 360
 319 GLPNMTOGFWENSMILDPGNVQKAVCHPTAMDGLGKDFRIIMCTKVTMDDELTAAHHEGH 378
 361 IYQDMAVAAOPFLIRNGANEGFHEAVGELMSLSAATPKHLKSIGLSPFOEDNETEINF 420
 379 IYQDMAVAAOPFLIRNGANEGFHEAVGELMSLSAATPKHLKSIGLSPFOEDNETEINF 438
 421 LKQALTIYGTLPFTYMLEKRWMTVFKGELIPKQDMKKWEMKREIVGVVEVPDDETYC 480
 439 LKQALTIYGTLPFTYMLEKRWMTVFKGELIPKQDMKKWEMKREIVGVVEVPDDETYC 498
 481 DPASLFHVSDYSFIRYRTITLYQFOFQALCOAKHNEBRLHKCDISNSTEAGOKL 536
 499 DPASLFHVSDYSFIRYRTITLYQFOFQALCOAKHNEBRLHKCDISNSTEAGOKL 554

RESULT 11
 US-10-121-049-72
 ; Sequence 72, Application US/10121049
 ; Publication No. US2003002239A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen
 APPLICANT: Deforge, Laura
 APPLICANT: Desnoyers, Luc
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerlitsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gunney, Austin L.
 APPLICANT: Sherwood, Steven
 APPLICANT: Smith, Victoria
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K
 APPLICANT: Wood, William
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P330R1C17
 CURRENT APPLICATION NUMBER: US/10/121,049
 CURRENT FILING DATE: 2002-04-12
 Prior Application removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 72
 LENGTH: 555
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-121-049-72

Query Match 74.9%; Score 2897; DB 9; Length 555;
 Best Local Similarity 99.8%; Pred. No. 5.3e-232;
 Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 STEEOATFLDKFHEAEDLFYOSLASMYNTNITEENVONNNAAGKMSAFLEKOST 60
 19 STEEOATFLDKFHEAEDLFYOSLASMYNTNITEENVONNNAAGKMSAFLEKOST 78
 61 LAQMTPLDQIOMLVKLOLQALQONGSSVLESDSKRLNTLNTMSTIYSTGVCNPNP 120
 79 LAQMTPLDQIOMLVKLOLQALQONGSSVLESDSKRLNTLNTMSTIYSTGVCNPNP 138
 121 QECILLEPGLNEIMANSIDYNERLMAWESRSEVSKOLRPLYEEVVLKEMARANYED 180
 139 QECILLEPGLNEIMANSIDYNERLMAWESRSEVSKOLRPLYEEVVLKEMARANYED 198
 181 YGDYWRGDEYVNGVGYDYSRGQLIEDVEHFEETIKPLYEHLHAYRAKLMNAPSYISP 240
 199 YGDYWRGDEYVNGVGYDYSRGQLIEDVEHFEETIKPLYEHLHAYRAKLMNAPSYISP 258
 241 IGLPRAHLIGDMWGRFWTNLSLVFPGOKPNIDVTDAVDAQADRIEFAEKFFVS 300
 259 IGLPRAHLIGDMWGRFWTNLSLVFPGOKPNIDVTDAVDAQADRIEFAEKFFVS 318
 301 GLPNMTOGFWENSMILDPGNVQKAVCHPTAMDGLGKDFRIIMCTKVTMDDELTAAHHEGH 360
 319 GLPNMTOGFWENSMILDPGNVQKAVCHPTAMDGLGKDFRIIMCTKVTMDDELTAAHHEGH 378
 361 IYQDMAVAAOPFLIRNGANEGFHEAVGELMSLSAATPKHLKSIGLSPFOEDNETEINF 420
 379 IYQDMAVAAOPFLIRNGANEGFHEAVGELMSLSAATPKHLKSIGLSPFOEDNETEINF 438
 421 LKQALTIYGTLPFTYMLEKRWMTVFKGELIPKQDMKKWEMKREIVGVVEVPDDETYC 480
 439 LKQALTIYGTLPFTYMLEKRWMTVFKGELIPKQDMKKWEMKREIVGVVEVPDDETYC 498
 481 DPASLFHVSDYSFIRYRTITLYQFOFQALCOAKHNEBRLHKCDISNSTEAGOKL 536
 499 DPASLFHVSDYSFIRYRTITLYQFOFQALCOAKHNEBRLHKCDISNSTEAGOKL 554

RESULT 12
 US-10-123-904-72
 ; Sequence 72, Application US/10123904

```
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C54
; CURRENT APPLICATION NUMBER: US/10/123,904
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 72
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-904-72
```

```
Query Match          74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEEQAKTFLEKFNHEADLFYQSSLASWNTNTITEENVQNNAGDKMSAFLEKQST 60
DB 19 STEEQAKTFLEKFNHEADLFYQSSLASWNTNTITEENVQNNAGDKMSAFLEKQST 78
QY 61 LAQMTPLQEIQNTLVKIQLOALQNGSSVLSSEKSKRLNTLNTSTIYSGKCNPNP 120
DB 79 LAQMTPLQEIQNTLVKIQLOALQNGSSVLSSEKSKRLNTLNTSTIYSGKCNPNP 138
QY 121 QECILLEPGLNEMANSIDYNERLWAMESWSEVQKQRLPYEEYVVLKNEMARANYED 180
DB 139 QECILLEPGLNEMANSIDYNERLWAMESWSEVQKQRLPYEEYVVLKNEMARANYED 198
QY 181 YGDYMRGDEYVNGVDYYSRGQLEDEVEHTEFEIKPLYEHLHAYYRAKLMAAPSYISP 240
DB 199 YGDYMRGDEYVNGVDYYSRGQLEDEVEHTEFEIKPLYEHLHAYYRAKLMAAPSYISP 258
QY 241 IGCPLPHLLGDMWGRFTWNLXSLTVPFGOKPNIDVTDAVDAQMRIFEKAEKFFVS 300
DB 259 IGCPLPHLLGDMWGRFTWNLXSLTVPFGOKPNIDVTDAVDAQMRIFEKAEKFFVS 318
QY 301 GLPNMTQGFWENSMITDGNVOKAVCHPTAMDLGKDFRILMCKRYTMDFLTAHHBMGH 360
DB 319 GLPNMTQGFWENSMITDGNVOKAVCHPTAMDLGKDFRILMCKRYTMDFLTAHHBMGH 378
QY 361 IYDMMAYAAQPFLLRNGANEGFHAAGVETMSLSAATPKHLKISIGLSPDQEDNETEINF 420
DB 379 IYDMMAYAAQPFLLRNGANEGFHAAGVETMSLSAATPKHLKISIGLSPDQEDNETEINF 438
QY 421 LKQALTIYGLTPEFTYMLEKRWMMVFKGELPKDQMKKWKMKREIYGVVEPVPHDETYC 480
DB 439 LKQALTIYGLTPEFTYMLEKRWMMVFKGELPKDQMKKWKMKREIYGVVEPVPHDETYC 498
QY 481 DPASLFHYSNDYSFIRYTYRTLYOFQFOALCOAAKHGEPILKHCIDISNSTEAGOKL 536
DB 499 DPASLFHYSNDYSFIRYTYRTLYOFQFOALCOAAKHGEPILKHCIDISNSTEAGOKL 554
```

RESULT 13

US-10-140-470-72

```
; Sequence 72, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 72
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-470-72
```

```
Query Match          74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEEQAKTFLEKFNHEADLFYQSSLASWNTNTITEENVQNNAGDKMSAFLEKQST 60
DB 19 STEEQAKTFLEKFNHEADLFYQSSLASWNTNTITEENVQNNAGDKMSAFLEKQST 78
QY 61 LAQMTPLQEIQNTLVKIQLOALQNGSSVLSSEKSKRLNTLNTSTIYSGKCNPNP 120
DB 79 LAQMTPLQEIQNTLVKIQLOALQNGSSVLSSEKSKRLNTLNTSTIYSGKCNPNP 138
QY 121 QECILLEPGLNEMANSIDYNERLWAMESWSEVQKQRLPYEEYVVLKNEMARANYED 180
DB 139 QECILLEPGLNEMANSIDYNERLWAMESWSEVQKQRLPYEEYVVLKNEMARANYED 198
QY 181 YGDYMRGDEYVNGVDYYSRGQLEDEVEHTEFEIKPLYEHLHAYYRAKLMAAPSYISP 240
DB 199 YGDYMRGDEYVNGVDYYSRGQLEDEVEHTEFEIKPLYEHLHAYYRAKLMAAPSYISP 258
QY 241 IGCPLPHLLGDMWGRFTWNLXSLTVPFGOKPNIDVTDAVDAQMRIFEKAEKFFVS 300
DB 259 IGCPLPHLLGDMWGRFTWNLXSLTVPFGOKPNIDVTDAVDAQMRIFEKAEKFFVS 318
QY 301 GLPNMTQGFWENSMITDGNVOKAVCHPTAMDLGKDFRILMCKRYTMDFLTAHHBMGH 360
DB 319 GLPNMTQGFWENSMITDGNVOKAVCHPTAMDLGKDFRILMCKRYTMDFLTAHHBMGH 378
QY 361 IYDMMAYAAQPFLLRNGANEGFHAAGVETMSLSAATPKHLKISIGLSPDQEDNETEINF 420
DB 379 IYDMMAYAAQPFLLRNGANEGFHAAGVETMSLSAATPKHLKISIGLSPDQEDNETEINF 438
QY 421 LKQALTIYGLTPEFTYMLEKRWMMVFKGELPKDQMKKWKMKREIYGVVEPVPHDETYC 480
DB 439 LKQALTIYGLTPEFTYMLEKRWMMVFKGELPKDQMKKWKMKREIYGVVEPVPHDETYC 498
QY 481 DPASLFHYSNDYSFIRYTYRTLYOFQFOALCOAAKHGEPILKHCIDISNSTEAGOKL 536
DB 499 DPASLFHYSNDYSFIRYTYRTLYOFQFOALCOAAKHGEPILKHCIDISNSTEAGOKL 554
```


DB 499 DPASLFHVSDDYSFIRYRTTLXQFOFQALCOAKHESGPHLKCDISNSTEAGQKL 554

RESULT 14

US-10-175-746-72
Sequence 72, Application US/10175746
Publication No. US20030027270A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Laureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C353
CURRENT APPLICATION NUMBER: US/10/175,746
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-746-72

Query Match

Best Local Similarity 99.8%; Score 2897; DB 9; Length 555;
Pred. No. 5.3e-232; Mismatches 0; Indels 0; Gaps 0;
Matches 535; Conservative 1;

QY 1 STEEQAKTFLDKFNHEADLFYQSSLASWNTNTITEENVQANNAGDKSAFLKEQST 60
DB 19 STEEQAKTFLDKFNHEADLFYQSSLASWNTNTITEENVQANNAGDKSAFLKEQST 78
QY 61 LAQWYPILOEIONLVKQLQALQONGSSVLSSEKSKRLNTILNTMSTIYSTGKYCNPDP 120
DB 79 LAQWYPILOEIONLVKQLQALQONGSSVLSSEKSKRLNTILNTMSTIYSTGKYCNPDP 138
QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKQRLPYEEYVVLKEMARANYED 180
DB 139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKQRLPYEEYVVLKEMARANYED 198
QY 181 YGDYWRDYEYNGVDYDSRGQLIEDVHTFEETIKPLYEHLAAVRAKLMNAPSYISP 240
DB 199 YGDYWRDYEYNGVDYDSRGQLIEDVHTFEETIKPLYEHLAAVRAKLMNAPSYISP 258
QY 241 IGCPLPAHLIGDMGFRWTNLSLTVPFGOKPNIIDVTDAVDAQMDAQRIFKEAEKFEVSV 300
DB 259 IGCPLPAHLIGDMGFRWTNLSLTVPFGOKPNIIDVTDAVDAQMDAQRIFKEAEKFEVSV 318
QY 301 GLPNMTQGFWENSMITDPGNVQKAVCHPTAWDLGKDFRILMCTKVTYMDDEFLTAHHEMGH 360
DB 319 GLPNMTQGFWENSMITDPGNVQKAVCHPTAWDLGKDFRILMCTKVTYMDDEFLTAHHEMGH 378
QY 361 IQYDMAVYAAQPFLLRNGANGFHEAVGEIMSLSAATPKHLKSGILSPDOEDNETEINF 420
DB 379 IQYDMAVYAAQPFLLRNGANGFHEAVGEIMSLSAATPKHLKSGILSPDOEDNETEINF 438
QY 421 LKQALFTVGTLPFTYMLEKRWMMVFKEGELPKDQMMKRWEMKREIYGVVPEVPHDETYC 480
DB 439 LKQALFTVGTLPFTYMLEKRWMMVFKEGELPKDQMMKRWEMKREIYGVVPEVPHDETYC 498

QY 481 DPASLFHVSDDYSFIRYRTTLXQFOFQALCOAKHESGPHLKCDISNSTEAGQKL 536
DB 499 DPASLFHVSDDYSFIRYRTTLXQFOFQALCOAKHESGPHLKCDISNSTEAGQKL 554

RESULT 15

US-10-176-918-72
Sequence 72, Application US/10176918
Publication No. US20030027275A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Laureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C382
CURRENT APPLICATION NUMBER: US/10/176,918
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-918-72

Query Match

Best Local Similarity 99.8%; Score 2897; DB 9; Length 555;
Pred. No. 5.3e-232; Mismatches 0; Indels 0; Gaps 0;
Matches 535; Conservative 1;

QY 1 STEEQAKTFLDKFNHEADLFYQSSLASWNTNTITEENVQANNAGDKSAFLKEQST 60
DB 19 STEEQAKTFLDKFNHEADLFYQSSLASWNTNTITEENVQANNAGDKSAFLKEQST 78
QY 61 LAQWYPILOEIONLVKQLQALQONGSSVLSSEKSKRLNTILNTMSTIYSTGKYCNPDP 120
DB 79 LAQWYPILOEIONLVKQLQALQONGSSVLSSEKSKRLNTILNTMSTIYSTGKYCNPDP 138
QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKQRLPYEEYVVLKEMARANYED 180
DB 139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVQKQRLPYEEYVVLKEMARANYED 198
QY 181 YGDYWRDYEYNGVDYDSRGQLIEDVHTFEETIKPLYEHLAAVRAKLMNAPSYISP 240
DB 199 YGDYWRDYEYNGVDYDSRGQLIEDVHTFEETIKPLYEHLAAVRAKLMNAPSYISP 258
QY 241 IGCPLPAHLIGDMGFRWTNLSLTVPFGOKPNIIDVTDAVDAQMDAQRIFKEAEKFEVSV 300
DB 259 IGCPLPAHLIGDMGFRWTNLSLTVPFGOKPNIIDVTDAVDAQMDAQRIFKEAEKFEVSV 318
QY 301 GLPNMTQGFWENSMITDPGNVQKAVCHPTAWDLGKDFRILMCTKVTYMDDEFLTAHHEMGH 360
DB 319 GLPNMTQGFWENSMITDPGNVQKAVCHPTAWDLGKDFRILMCTKVTYMDDEFLTAHHEMGH 378
QY 361 IQYDMAVYAAQPFLLRNGANGFHEAVGEIMSLSAATPKHLKSGILSPDOEDNETEINF 420
DB 379 IQYDMAVYAAQPFLLRNGANGFHEAVGEIMSLSAATPKHLKSGILSPDOEDNETEINF 438

QY 421 LKQALTYGTLPTTYMLEKRMWVFKGEIPKQDMKKMKMKREIVGVVEPHDETC 480
DB 439 LKQALTYGTLPTTYMLEKRMWVFKGEIPKQDMKKMKMKREIVGVVEPHDETC 498
QY 481 DPASLFHVSNDYSFIRYTRTLTQFOFQALCOAANKHGBPLHKDISNSTEAGOKL 536
DB 499 DPASLFHVSNDYSFIRYTRTLTQFOFQALCOAANKHGBPLHKDISNSTEAGOKL 554

RESULT 16

US-10-176-921-72
Sequence 72, Application US/10176921
Publication No. US20030027276A1

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Laureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C288
CURRENT APPLICATION NUMBER: US/10/176,921
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-921-72

Query Match 74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIEQAKTFLDKFNHEADLFYOSLSASWNTNTITEENVOANNAGDKWSAFLKEQST 60
DB 19 STIEQAKTFLDKFNHEADLFYOSLSASWNTNTITEENVOANNAGDKWSAFLKEQST 78
QY 61 LAOMYPLQEIOTLVKIQALQOQSSSVLSEDSKRLNTILNTMSTIYSTGKVCNDNP 120
DB 79 LAOMYPLQEIOTLVKIQALQOQSSSVLSEDSKRLNTILNTMSTIYSTGKVCNDNP 138
QY 121 QECILLEPGLNEIMANSIDYNERLWAMESWSEVGKQRLPYEEYVVLKNEAARAHYED 180
DB 139 QECILLEPGLNEIMANSIDYNERLWAMESWSEVGKQRLPYEEYVVLKNEAARAHYED 198
QY 181 YGDYWRGDEYVNGVDGYDSRGQLLEDEYHFEETIKPLYEHLAAVYAKLMAVPSYIS 240
DB 199 YGDYWRGDEYVNGVDGYDSRGQLLEDEYHFEETIKPLYEHLAAVYAKLMAVPSYIS 258
QY 241 IGCPLPAHLGDMWGRFTNLTSLVPGQKPNIDVTDAVDAQDAORIFKEAEKFFVS 300
DB 259 IGCPLPAHLGDMWGRFTNLTSLVPGQKPNIDVTDAVDAQDAORIFKEAEKFFVS 318
QY 301 GLPNTQGFWENSMLTDGNNVOKAVCHPTAMDIGKGFRLIMCTKYTMDDFLTAHHMGH 360
DB 319 GLPNTQGFWENSMLTDGNNVOKAVCHPTAMDIGKGFRLIMCTKYTMDDFLTAHHMGH 378
QY 361 IGYDAVYAAQPLLRNGANEGHEAVGEIMSLISAATPKHLASIGLLSPDFEDNETEINF 420

DB 379 IGYDAVYAAQPLLRNGANEGHEAVGEIMSLISAATPKHLASIGLLSPDFEDNETEINF 438
QY 421 LKQALTYGTLPTTYMLEKRMWVFKGEIPKQDMKKMKMKREIVGVVEPHDETC 480
DB 439 LKQALTYGTLPTTYMLEKRMWVFKGEIPKQDMKKMKMKREIVGVVEPHDETC 498
QY 481 DPASLFHVSNDYSFIRYTRTLTQFOFQALCOAANKHGBPLHKDISNSTEAGOKL 536
DB 499 DPASLFHVSNDYSFIRYTRTLTQFOFQALCOAANKHGBPLHKDISNSTEAGOKL 554

RESULT 17

US-10-137-865-72
Sequence 72, Application US/10137865
Publication No. US20030032155A1

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Laureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C154
CURRENT APPLICATION NUMBER: US/10/137,865
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-137-865-72

Query Match 74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIEQAKTFLDKFNHEADLFYOSLSASWNTNTITEENVOANNAGDKWSAFLKEQST 60
DB 19 STIEQAKTFLDKFNHEADLFYOSLSASWNTNTITEENVOANNAGDKWSAFLKEQST 78
QY 61 LAOMYPLQEIOTLVKIQALQOQSSSVLSEDSKRLNTILNTMSTIYSTGKVCNDNP 120
DB 79 LAOMYPLQEIOTLVKIQALQOQSSSVLSEDSKRLNTILNTMSTIYSTGKVCNDNP 138
QY 121 QECILLEPGLNEIMANSIDYNERLWAMESWSEVGKQRLPYEEYVVLKNEAARAHYED 180
DB 139 QECILLEPGLNEIMANSIDYNERLWAMESWSEVGKQRLPYEEYVVLKNEAARAHYED 198
QY 181 YGDYWRGDEYVNGVDGYDSRGQLLEDEYHFEETIKPLYEHLAAVYAKLMAVPSYIS 240
DB 199 YGDYWRGDEYVNGVDGYDSRGQLLEDEYHFEETIKPLYEHLAAVYAKLMAVPSYIS 258
QY 241 IGCPLPAHLGDMWGRFTNLTSLVPGQKPNIDVTDAVDAQDAORIFKEAEKFFVS 300
DB 259 IGCPLPAHLGDMWGRFTNLTSLVPGQKPNIDVTDAVDAQDAORIFKEAEKFFVS 318
QY 301 GLPNTQGFWENSMLTDGNNVOKAVCHPTAMDIGKGFRLIMCTKYTMDDFLTAHHMGH 360

Db 319 GLPNMTQGFENSMULTDPCGNVOKAVCHPTAMDIGKGFRIIMCTKYTMDFLTAHHKMGH 378
QY 361 IQYDMAVAAOPELLRNGANGEGHEAVEIGMSLSAATPKHLKSIIGLSPDFQEDNETEINF 420
Db 379 IQYDMAVAAOPELLRNGANGEGHEAVEIGMSLSAATPKHLKSIIGLSPDFQEDNETEINF 438
QY 421 LKQALITVGLPTPTMYLKKRMWVFKEGELPKQOMKMKWEMKREIYGVEVPHPDETTC 480
Db 439 LKQALITVGLPTPTMYLKKRMWVFKEGELPKQOMKMKWEMKREIYGVEVPHPDETTC 498
QY 481 DPASLFHVSNDYSFIRYTRTLVYQFOFQALCOAAKHGPHLKCIDSNSTEAGOKL 536
Db 499 DPASLFHVSNDYSFIRYTRTLVYQFOFQALCOAAKHGPHLKCIDSNSTEAGOKL 554

RESULT 18
US-10-474-72

Sequence 72, Application US/10140474
Publication No. US20030032156A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Laureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey J.
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33081C162
CURRENT APPLICATION NUMBER: US/10/140,474
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-474-72

Query Match 74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 STIEQAKTFLDKFNHAEADLFYQSSLSASWNTNTITEENVQNMNAGDKWSAFLEKOST 60
Db 19 STIEQAKTFLDKFNHAEADLFYQSSLSASWNTNTITEENVQNMNAGDKWSAFLEKOST 78
QY 61 LAQYPLQEIQNLTVKQLQALQONGSSVLSSEDKSKRLNTILNTMSTIYSTGKYCNPDP 120
Db 79 LAQYPLQEIQNLTVKQLQALQONGSSVLSSEDKSKRLNTILNTMSTIYSTGKYCNPDP 138
QY 121 QECILLEPGLNEIMANSIDYNERLWAMESWSEVKGOLRPLYEYVVLKNEKARAHYED 180
Db 139 QECILLEPGLNEIMANSIDYNERLWAMESWSEVKGOLRPLYEYVVLKNEKARAHYED 198
QY 181 YGDYWRGDEYVNGVDYDSRGQLIEDVHTFEETIKPLYEHLHAYVRAKLMAVPSYISP 240
Db 199 YGDYWRGDEYVNGVDYDSRGQLIEDVHTFEETIKPLYEHLHAYVRAKLMAVPSYISP 258
QY 241 IGCPLAHLGDMGWFNTNLSLTVPGQKPNIDVTAMDVAQDAORIFKEAKFEFVS 300
Db 259 IGCPLAHLGDMGWFNTNLSLTVPGQKPNIDVTAMDVAQDAORIFKEAKFEFVS 318

QY 301 GLPNMTQGFENSMULTDPCGNVOKAVCHPTAMDIGKGFRIIMCTKYTMDFLTAHHKMGH 360
Db 319 GLPNMTQGFENSMULTDPCGNVOKAVCHPTAMDIGKGFRIIMCTKYTMDFLTAHHKMGH 378
QY 361 IQYDMAVAAOPELLRNGANGEGHEAVEIGMSLSAATPKHLKSIIGLSPDFQEDNETEINF 420
Db 379 IQYDMAVAAOPELLRNGANGEGHEAVEIGMSLSAATPKHLKSIIGLSPDFQEDNETEINF 438
QY 421 LKQALITVGLPTPTMYLKKRMWVFKEGELPKQOMKMKWEMKREIYGVEVPHPDETTC 480
Db 439 LKQALITVGLPTPTMYLKKRMWVFKEGELPKQOMKMKWEMKREIYGVEVPHPDETTC 498
QY 481 DPASLFHVSNDYSFIRYTRTLVYQFOFQALCOAAKHGPHLKCIDSNSTEAGOKL 536
Db 499 DPASLFHVSNDYSFIRYTRTLVYQFOFQALCOAAKHGPHLKCIDSNSTEAGOKL 554

RESULT 19
US-10-142-431-72

Sequence 72, Application US/10142431
Publication No. US20030036179A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Laureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey J.
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33081C251
CURRENT APPLICATION NUMBER: US/10/142,431
CURRENT FILING DATE: 2002-05-10
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-142-431-72

Query Match 74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 STIEQAKTFLDKFNHAEADLFYQSSLSASWNTNTITEENVQNMNAGDKWSAFLEKOST 60
Db 19 STIEQAKTFLDKFNHAEADLFYQSSLSASWNTNTITEENVQNMNAGDKWSAFLEKOST 78
QY 61 LAQYPLQEIQNLTVKQLQALQONGSSVLSSEDKSKRLNTILNTMSTIYSTGKYCNPDP 120
Db 79 LAQYPLQEIQNLTVKQLQALQONGSSVLSSEDKSKRLNTILNTMSTIYSTGKYCNPDP 138
QY 121 QECILLEPGLNEIMANSIDYNERLWAMESWSEVKGOLRPLYEYVVLKNEKARAHYED 180
Db 139 QECILLEPGLNEIMANSIDYNERLWAMESWSEVKGOLRPLYEYVVLKNEKARAHYED 198
QY 181 YGDYWRGDEYVNGVDYDSRGQLIEDVHTFEETIKPLYEHLHAYVRAKLMAVPSYISP 240
Db 199 YGDYWRGDEYVNGVDYDSRGQLIEDVHTFEETIKPLYEHLHAYVRAKLMAVPSYISP 258

QY 241 IGCPLAHLLGDMGFRFTNLXSLVPEFGOKPNIDVTAMVDQAMDARIFKEAEKFEVS 300
DB 259 IGCPLAHLLGDMGFRFTNLXSLVPEFGOKPNIDVTAMVDQAMDARIFKEAEKFEVS 318
QY 301 GLPMTQGFWMNSMLTDPGNVOKAVCHPTAMDLGKDFRLIMCTKYTMDDFLTAHHEMGH 360
DB 319 GLPMTQGFWMNSMLTDPGNVOKAVCHPTAMDLGKDFRLIMCTKYTMDDFLTAHHEMGH 378
QY 361 IGYDAAVAAOPFLLRNGANGFHEAVGEIMSLAATPKHLKSLGILSPDFOEDENETEINF 420
DB 379 IGYDAAVAAOPFLLRNGANGFHEAVGEIMSLAATPKHLKSLGILSPDFOEDENETEINF 438
QY 421 LKQALITVGLTPEFTYMLEKRMWVFKGEIPKQDMKKWMEKREIVGVVEPVDHETYC 480
DB 439 LKQALITVGLTPEFTYMLEKRMWVFKGEIPKQDMKKWMEKREIVGVVEPVDHETYC 498
QY 481 DPASLFHVSNDYSFIRYRTLYQFOFQALCOAKHGGPLHKCDISNSTEAGOKL 536
DB 499 DPASLFHVSNDYSFIRYRTLYQFOFQALCOAKHGGPLHKCDISNSTEAGOKL 554

RESULT 20

US-10-143-114-72
Sequence 72, Application US/10143114
Publication No. US20030036180A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330301C211
CURRENT APPLICATION NUMBER: US/10/143,114
CURRENT FILING DATE: 2002-05-09
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-143-114-72

Query Match

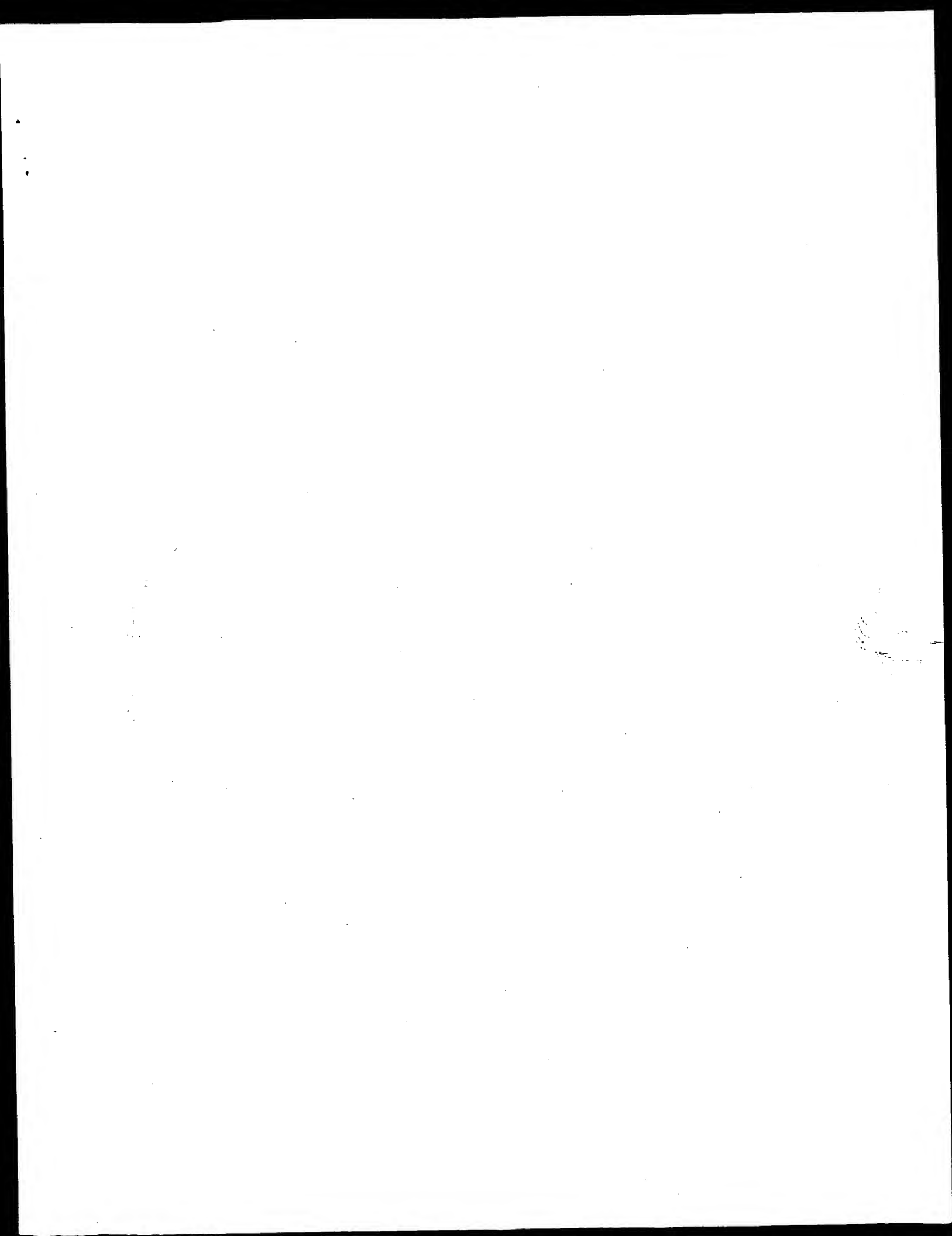
Best Local Similarity 74.9%; Score 2897; DB 9; Length 555;
Pred. No. 5.3e-232;

Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIEBQAKFTLDKFNHEADLFYQSSLSAMNTNITEENVONMANNAGDKMSAPLKEQST 60
DB 19 STIEBQAKFTLDKFNHEADLFYQSSLSAMNTNITEENVONMANNAGDKMSAPLKEQST 78
QY 61 LAQMTPLQETIQLVYKLOLQALQONGSSVLSSEKSKRLNTILNTMSTIYSTGKVCNPDP 120
DB 79 LAQMTPLQETIQLVYKLOLQALQONGSSVLSSEKSKRLNTILNTMSTIYSTGKVCNPDP 138
QY 121 QECILLEPGLNEIVANSIDYNERLMAWESRSEVGKQRLPYEEYVVLKNEMARAHYED 180
DB 139 QECILLEPGLNEIVANSIDYNERLMAWESRSEVGKQRLPYEEYVVLKNEMARAHYED 198
QY 181 YGDYWRGDEYVNGVDYDYSRQGLIEDVEHTEFEIKPLYEHLHAVVRAKLMNAVPSYISP 240

DB 199 YGDYWRGDEYVNGVDYDYSRQGLIEDVEHTEFEIKPLYEHLHAVVRAKLMNAVPSYISP 258
QY 241 IGCPLAHLLGDMGFRFTNLXSLVPEFGOKPNIDVTAMVDQAMDARIFKEAEKFEVS 300
DB 259 IGCPLAHLLGDMGFRFTNLXSLVPEFGOKPNIDVTAMVDQAMDARIFKEAEKFEVS 318
QY 301 GLPMTQGFWMNSMLTDPGNVOKAVCHPTAMDLGKDFRLIMCTKYTMDDFLTAHHEMGH 360
DB 319 GLPMTQGFWMNSMLTDPGNVOKAVCHPTAMDLGKDFRLIMCTKYTMDDFLTAHHEMGH 378
QY 361 IGYDAAVAAOPFLLRNGANGFHEAVGEIMSLAATPKHLKSLGILSPDFOEDENETEINF 420
DB 379 IGYDAAVAAOPFLLRNGANGFHEAVGEIMSLAATPKHLKSLGILSPDFOEDENETEINF 438
QY 421 LKQALITVGLTPEFTYMLEKRMWVFKGEIPKQDMKKWMEKREIVGVVEPVDHETYC 480
DB 439 LKQALITVGLTPEFTYMLEKRMWVFKGEIPKQDMKKWMEKREIVGVVEPVDHETYC 498
QY 481 DPASLFHVSNDYSFIRYRTLYQFOFQALCOAKHGGPLHKCDISNSTEAGOKL 536
DB 499 DPASLFHVSNDYSFIRYRTLYQFOFQALCOAKHGGPLHKCDISNSTEAGOKL 554

Search completed: May 26, 2003, 17:59:11
Job time : 41 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 17:42:18 ; Search time 37 Seconds

(without alignments)
1870.724 Million cell updates/sec

Title: US-09-978-385-2_COPY_19_738

Sequence: 1 STIEBOAKTFIDKFHEAD.....DNSLEFLGIQPTLGPPNPP 720

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR:73:*
2: PIR:*
3: PIR:*
4: PIR:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3869	100.0	804	2	T14762
2	1337	34.6	732	1	S05238
3	1337	34.6	1306	1	A31759
4	1334	34.5	732	1	A35655
5	1334	34.5	1312	1	A34171
6	1312	33.9	1193	2	JC2489
7	1310	33.9	1313	1	JC2038
8	1283	33.2	737	1	A34402
9	1283	33.2	1309	1	S35484
10	1054.5	27.3	611	2	S65472
11	1027.5	26.6	630	2	JC5374
12	1022	26.4	615	2	A57532
13	638.5	16.5	907	2	T15792
14	157	4.1	532	2	C83696
15	154	4.0	502	2	AF1310
16	147	3.8	502	2	AE1682
17	139.5	3.6	987	2	AI2011
18	139	3.6	987	2	B82938
19	136	3.5	611	2	D82881
20	135	3.5	611	2	D69943
21	125	3.2	627	1	D40048
22	124.5	3.2	987	2	I48373
23	124	3.2	538	2	E72561
24	122	3.2	642	2	E98000
25	121	3.1	607	2	AB3511
26	121	3.1	1034	2	T30574
27	119	3.0	990	2	S23416
28	118	3.0	642	2	G95129
29	117.5	3.0	1339	2	A84683

Hit No. 1/5
too late

30	116	3.0	963	2	C90535	conserved hypothet
31	115.5	3.0	611	2	A75573	probable oligendo
32	115.5	3.0	3433	1	S28381	utrophin - human
33	115	3.0	950	2	A71655	hypothetical prote
34	115	3.0	1780	2	T17272	hypothetical prote
35	114.5	2.9	3856	2	T51174	ataxia-telangiecta
36	114	2.9	685	2	F75370	trypsin-like protea
37	113.5	2.9	3655	2	T38084	TRAP-like protein
38	113	2.9	772	2	A10968	probable glycosyl
39	113	2.9	902	2	E90270	conserved hypothet
40	113	2.9	1575	2	G82905	conserved hypothet
41	113	2.9	6669	2	S55024	nebulin, skeletal
42	112.5	2.9	808	2	A82877	hypothetical prote
43	112	2.9	524	2	B82202	thermostable carbo
44	111.5	2.9	282	2	H97226	protein containing
45	111	2.9	515	2	C71158	probable thermosta

ALIGNMENTS

RESULT 1

T14762
hypothetical protein DKFZp434A014.1 - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C/Accession: T14762

R/Wambuit, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, August 1999

A/Reference number: Z18181

A/Accession: T14762

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-804 <RAW>

A/Cross-references: EMBL:AL110224

A/Experimental source: adult testis; clone DKFZp434A014

C/Genetics:

A/Note: DKFZp434A014.1

Query Match	100.0%	Score 3869	DB 2	Length 804
Best Local Similarity	100.0%	Pred. No. 3,5e-265		
Matches 720	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	STIEBOAKTFIDKFHEADELFYQSSLASMNNTNITEENYQNNMAGDKMSAFLEKST	60	
DB	18	STIEBOAKTFIDKFHEADELFYQSSLASMNNTNITEENYQNNMAGDKMSAFLEKST	77	
QY	61	IAQMTPLQETQNLTKLQALQNGSSVLSDEKSKRLNTLNTMSTYSGKYNCPNP	120	
DB	78	IAQMTPLQETQNLTKLQALQNGSSVLSDEKSKRLNTLNTMSTYSGKYNCPNP	137	
QY	121	QECLELLEPLNITMANSIDYNERLWMSWSEVQKRLPYEEVYVTKNMAANHYED	180	
DB	138	QECLELLEPLNITMANSIDYNERLWMSWSEVQKRLPYEEVYVTKNMAANHYED	197	
QY	181	YGDYWRGDEYVNGVDYDYSRGQILEVEHTEFEIKPLYEHLHAYRAKLNNAVYSISP	240	
DB	198	YGDYWRGDEYVNGVDYDYSRGQILEVEHTEFEIKPLYEHLHAYRAKLNNAVYSISP	257	
QY	241	IGCLPAHLILGDMGFWNTNLSLYVPGCKPNIDVTAMDAMDADQRIFEAKRFVSV	300	
DB	258	IGCLPAHLILGDMGFWNTNLSLYVPGCKPNIDVTAMDAMDADQRIFEAKRFVSV	317	
QY	301	GLPNTQGFWMSMLTDPGNYQKAVCHPTAMDILGDFRILMCTKVTMDDELTAHHNGH	360	
DB	318	GLPNTQGFWMSMLTDPGNYQKAVCHPTAMDILGDFRILMCTKVTMDDELTAHHNGH	377	
QY	361	IQYDVAAYAAPFLNKGANGEGFHEAVGEIMSLSATPKHLKSGILSPDEQEDNETINF	420	
DB	378	IQYDVAAYAAPFLNKGANGEGFHEAVGEIMSLSATPKHLKSGILSPDEQEDNETINF	437	
QY	421	LLKQALITVGLPFLPYHLEKRWAVFGEIIRKQDMKRWEMKREIYGVPEVDHETTC	480	

Db 438 LKQALITVGTLPETYLEKRWMMVFKEGELIPKQDMKRWMEKREIVGVPEVHDETYC 497
 QY 481 DPASLFHNSNDYSPFRYRTPLTQFOFQALCOAKHGGPLHKCDISNSTEAGOKLNNML 540
 Db 498 DPASLFHNSNDYSPFRYRTPLTQFOFQALCOAKHGGPLHKCDISNSTEAGOKLNNML 557
 QY 541 RLKSEPTWTLLENVYGAKNMNVRLNTEPEPLTWLKDQKNSFVGMSTDMSPYADQSI 600
 Db 558 RLKSEPTWTLLENVYGAKNMNVRLNTEPEPLTWLKDQKNSFVGMSTDMSPYADQSI 617
 QY 601 KVRISLKSALDDKAYEMNDNEMYLFRSSVAAYAMQYLLKYNNQMLLGEEDVRYANLKKR 660
 Db 618 KVRISLKSALDDKAYEMNDNEMYLFRSSVAAYAMQYLLKYNNQMLLGEEDVRYANLKKR 677
 QY 661 ISFNFTAPKPNVSDIIPTEVEKAIKRSRIRINDAFELNDNSLEFLGIQPTLGPQPP 720
 Db 678 ISFNFTAPKPNVSDIIPTEVEKAIKRSRIRINDAFELNDNSLEFLGIQPTLGPQPP 737

RESULT 2

peptidyl-dipeptidase A (EC 3.4.15.1) precursor, testicular splice form - human
 A:Accession: S05238
 A:Residues: 1-732 <LAT>
 A:Cross-references: EMBL:X16295; NID:q28264; PIDN:CAA3362.1; PID:q28265
 R:Ehlers, M.R.W.; Fox, E.A.; Strydom, D.J.; Riordan, J.F.
 Proc. Natl. Acad. Sci. U.S.A. 86, 7741-7745, 1989
 A:Title: Molecular cloning of human testicular angiotensin-converting enzyme: the testis
 A:Reference number: A33979; MUID:90046671; PMID:2554286
 A:Accession: A33979
 A:Molecule type: mRNA
 A:Residues: 1-732 <ENL>
 A:Cross-references: GB:M26657; NID:q338666; PIDN:AAA60611.1; PID:q338667
 A:Experimental source: clones R1.2 and T88
 A:Note: neither the complete nucleic acid sequence nor the complete translation are shown
 C:Comment: For the renal and pulmonary splice form, see PIR:A51759.
 C:Genetics:
 A:Gene: GDB:DCPI; ACE
 A:Cross-references: GDB:119840; OMIM:106180
 A:Map position: 17q23-17q23
 C:Function:
 A:Description: catalyzes the hydrolysis of dipeptides from the carboxyl end of polypeptide
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; glycoprotein; metalloproteinase; peptidyl-dipeptide hydrolase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-732/Product: peptidyl dipeptidase I #status predicted <MAT>
 F:686-702/Domain: transmembrane #status predicted <RM>
 F:103-121,140,186,368,617,651/Binding site: carbonylate (Asn) (covalent) #status predicted
 F:414,418,434/Binding site: zinc, catalytic (His, His, Glu) #status predicted
 F:415/Active site: Glu #status predicted

Query Match 34.6% Score 1337; DB 1; Length 732;
 Best Local Similarity 41.7% Pred. No.2.4e+86;
 Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

QY 2 TIEQAKTFLDKFNHAEADLFYOSLSAMNTNTITE-----NYQNNNMGDKWSA 53
 Db 70 TDEAASKFVEYEDTQSVVWNEVEANMNTNTITESTILKQNNHNT----- 123
 QY 54 FLKEQSTLAQMTPLQFIONLVKIQLOALQONGSSVLESDSKRLNTILNMTSTYSGK 113
 Db 124 --LKYGTOARKEDVQALQNTYIKRIKRVQDLERALPAQGLEENKTLIDMETTYSVAT 181
 QY 114 VCNPNPQECULLLEGLMIANSLDYNERLWAMSWRSSEYKQRLPYEEYVVLKNNMA 173

Db 182 VCHPNG--SCIOLEPDLTVMAATSRKYEDELLWAMEBMRKAGRALIOEPKYEVELINQAA 239
 QY 174 RANHYEDGDYWRGDEYVNGCYDYSRQGLIEDVEHFEETIKPLYEHHAAYRAKLMA 233
 Db 240 RLNGYVDAGDSMSMETPSLE-----QDERLLEQELPLNLHAAYRALHRH 289
 QY 234 Y-PSYISPIGCLPAHLIDGMGRFWNTLSLYVPEGQKPNIDVTAMVQDAQRIKRE 292
 Db 290 YGAQHINLEGPYIAHLIDGMGRFWNTLSLYVPEGQKPNIDVTAMVQDAQRIKRE 349
 QY 293 AEFYFVSVGLPNNQGFENSMITDGNQKAVCHETANDLKG-DEFILCTKYTMDF 351
 Db 350 ADEFYFVSVGLPNNQGFENSMITDGNQKAVCHETANDLKG-DEFILCTKYTMDF 409
 QY 352 LTAHHEGHIQYDMAFAAPPELLRNGANGFPAVEGSSAATPCKHLSIGLSLPDQ 411
 Db 410 VYAHHEGHIQYDMAFAAPPELLRNGANGFPAVEGSSAATPCKHLSIGLSLPDQ 469
 QY 412 EDNETEINFLKQALITVGTLPETYLEKRWMMVFKEGELIPKQDMKRWMEKREIVGVPE 471
 Db 470 SD-ENHINFLKQALITVGTLPETYLEKRWMMVFKEGELIPKQDMKRWMEKREIVGVPE 528
 QY 472 PVHDETYCDPASFLEHNSNDYSPFRYRTPLTQFOFQALCOAKHGGPLHKCDISNSTE 531
 Db 529 PVHDETYCDPASFLEHNSNDYSPFRYRTPLTQFOFQALCOAKHGGPLHKCDISNSTE 588
 QY 532 AGQRLNATAMKLGSRWPEAMQITGPQNNASAMLSYKPLDMLRTNELHGEKLGMP 588
 Db 589 AGQRLNATAMKLGSRWPEAMQITGPQNNASAMLSYKPLDMLRTNELHGEKLGMP 648
 QY 589 STDMSPYADQSI 599
 Db 649 QYMTWPSARS 659

RESULT 3

peptidyl-dipeptidase A (EC 3.4.15.1) precursor, renal and pulmonary splice form - human
 A:Accession: A31759
 A:Residues: 1-1306 <SD>
 A:Cross-references: GB:U04144; NID:q178285; PIDN:AAA51684.1; PID:q178286
 A:Experimental source: kidney
 A:Note: parts of this sequence, including the amino end of the mature protein, were derived from the cDNA of the human angiotensin-converting enzyme (ACE); CD143; dipeptidyl carboxypeptidase
 R:Takeuchi, K.; Shimizu, T.; Ohishi, N.; Seyama, Y.; Takaku, F.; Yotsumoto, H.
 J. Biochem. 106, 442-445, 1989
 A:Title: Purification of human lung angiotensin-converting enzyme by high-performance
 A:Reference number: P00004; MUID:90110025; PMID:2558109
 A:Accession: P00004
 A:Molecule type: protein
 A:Residues: 'XX', '32-34', 'E', '36-37', 'X', '39-41', 'R', '43-46' <TAK>
 A:Experimental source: lung
 C:Comment: This splice form is found in many tissues, in particular kidney and lung v
 C:Genetics:
 A:Gene: GDB:DCPI; ACE
 A:Cross-references: GDB:119840; OMIM:106180
 A:Map position: 17q23-17q23
 C:Function:
 A:Description: catalyzes the hydrolysis of dipeptides from the carboxyl end of polypeptide
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; blood pressure control; glycoprotein; kidney; lung;
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-1306/Product: peptidyl dipeptidase I #status predicted <MAT>

C: Superfamily: mammalian peptidyl-di-peptidase A
C: Keywords: alternative splicing; peptidyl dipeptide hydrolase; transmembrane p

Title: Partial protein sequence of mouse and bovine kidney angiotensin convert
 Reference number: A61477; MUID:88215372; PMID:2835538

A:Accession: A61477
 A:Status: Preliminary
 A:Molecule type: protein
 A:Residues: 35-54 <BB3>
 A:Experimental source: kidney
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; blood pressure control; membrane protein; peptidyl-dipeptidase I
 F:1-34/Domain: signal sequence #status predicted <SIG>
 F:35-1312/Product: peptidyl dipeptidase I #status predicted <MAT>

Query Match 34.5%; Score 1334; DB 1; Length 1312;
 Best Local Similarity 42.6%; Pred. No. 9.6e-86;
 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

QY 2 TTEQAKTFLEKHEAEADLFYQSSLSMNNTNTEENQNMNAGDKMSAFLEKSTL 61
 DB TDEKADRFVEYRQAVLNLEYAEANWQNTNITTEGSKITLKEKTEVSNHLKCTR 708
 QY 62 AQMTPLAEIQNTLVKLOLQNGSSVLSSEKSKRLNTLNTSTYTGKVCNPNPQ 121
 DB AKTQDVSNFQSSIKRIKQNDRAVLPEKLEEVNQILDMETTYLSNICYTNG-- 766
 QY 122 ECLLEPGELNMANSLDYNERLWMSSEVQKOLRPLEEYVVLKNEKARAHEDY 181
 DB TCMLEPDLTMMATSRKYEELLMAMKSWDRKVRALLPEPKYVEFSNKIAKNGYTDA 826
 QY 182 GDYMRGDEVNGVDGYDSRGQLLEDVHEETEEIKPLYEHLHAYRAKLMNAYS-YISP 240
 DB GDSWRSLEYEDNLE-----QDLKLYQELQPLYLNLHAYRSLHNGSEYINL 876
 QY 241 IGCLEPAHLIGDMGSEFTNLVSLVPEGQKNIDVTDAVDAQDAQRIFEAKKFPVS 300
 DB DGPPIAHLIGNMMAQSWNITDYLVAPEPSAPINIDATEAMIKQWTPRIFEKANPFTSL 936
 QY 301 GLPNTQGFEMNSMLTDGYNOKACHPDANDLIG--DEFTLMCTKVTMDPLFAHHEM 359
 DB 877 DGPPIAHLIGNMMAQSWNITDYLVAPEPSAPINIDATEAMIKQWTPRIFEKANPFTSL 936
 QY 937 GLPPEPFEMNSMLEKPTDGEVYVCHSANDFYNGKDFRIKQCTSVAMDELVAHHEM 996
 DB 360 HIOYDMAYAAQPFLLRNGANGFHEAVGEIWSLAPPRKILSGILSPDFQEDNEFEIN 419
 QY 997 HIOYDMAYAAQPFLLRNGANGFHEAVGEIWSLAPPRKILSGILSPDFQEDNEFEIN 419
 DB 420 FLTKALALTYGLPTVYLMKEMRMVVEGELPKQOMKMKMKREIYGVVEPYPHDET 479
 QY 1056 FLTKALALTYGLPTVYLMKEMRMVVEGELPKQOMKMKMKREIYGVVEPYPHDET 479
 DB 480 CDPASLFFHSNDYSFIRYRTTLYQFOFQALCOAKHKGEPHLCIDISNSTEAGOKLFNM 539
 QY 1116 PDPSGKFNHPANVPYRFRVSIIOFQHEALCRAAGTGPHKCDIYQSEKAGLADA 1175
 DB 540 LRLGSEPTWTLAENVGAKNNVRLNTEPELFTWL-KDONK--NSFVGM-STDMSP 594
 QY 1176 KRLGSKPWPPEAMKLTICQPNMSASAMNVPKPLEMTLVTEENRRGFTLIGPEYMAP 1233
 DB

RESULT 6
 JC2489
 peptidyl-dipeptidase A (EC 3.4.15.1) - chicken

N:Alternate names: angiotensin converting enzyme
 C:Species: Gallus gallus (chicken)
 C:Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 28-May-1999
 C:Accession: JC2489
 R:Author: C.R.; Thomas Jr., K.E.; Bernstein, K.E.
 Biochem. Biophys. Res. Commun. 205, 1916-1921, 1994
 A:Title: Chicken lacks the testis specific isoform of angiotensin converting enzyme four
 A:Reference number: JC2489; MUID:95110342; PMID:7811282
 A:Accession: JC2489
 A:Molecule type: mRNA
 A:Residues: 1-1193 <EST>
 A:Cross-references: GB:U40175; NID:9685168; PIDN:AAA75554.1; PID:9994708
 C:Comment: This enzyme is a zinc dependant dipeptidyl carboxypeptidase that cleaves a va
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: metal binding; peptidyl-dipeptide hydrolase; zinc

F:316,331,914,929/Binding site: zinc, catalytic (Glu, His, Glu, His) #status predicted

Query Match 33.9%; Score 1312; DB 2; Length 1193;
 Best Local Similarity 40.4%; Pred. No. 3e-84;
 Matches 264; Conservative 110; Mismatches 243; Indels 36; Gaps 10;

QY 4 EEOAKTFLEKHEAEADLFYQSSLSMNNTNTEENQNMNAGDKMSAFLEKSTL 63
 DB EAOAKTFLESEYNSTAEVNNAYTEASWEYNTNITDHKNKVELEKMLAMSKHTIEGMBAR 603
 QY 64 MYPLAEIQNTLVKLOLQNGSSVLSSEKSKRLNTLNTSTYTGKVCNPNP--P 120
 DB 604 QEPSPQFQETVYRIKLSVTERALPDELEKETVTLSDMETTYSAKVCRENNTNHP 663
 QY 121 QECLEPGELNMANSLDYNERLWMSSEVQKOLRPLEEYVVLKNEKARAHEDY 180
 DB 664 -----LDPLDTLARSROYNELFPMKQMWASGAKIKDKRYVELSNKAALVNGYTD 718
 QY 181 YGDMRGDEVNGVDGYDSRGQLLEDVHEETEEIKPLYEHLHAYRAKLMNAY-PSYIS 239
 DB 719 NGAYMRSLEYETPFE-----EDLERLYQLQPLYLNLHAYRRALYKNGAGEHIS 768
 QY 240 PICGLPAHLIGDMGSEFTNLVSLVPEGQKNIDVTDAVDAQDAQRIFEAKKFPVS 299
 DB 769 LKGPPIAHLIGNMMAQSWNITDYLVAPEPSAPINIDATEAMIKQWTPRIFEKANPFTSL 936
 QY 300 VGLPNTQGFEMNSMLTDGYNOKACHPDANDL--GKGFRLMCTKVTMDPLFAHHEM 358
 DB 829 LGLIPDQEFWDMKSMLEKPADGREVYCHASAMDFYRKDFRIKQCTSVAMDELVAHHEM 888
 QY 359 GHIOYDMAYAAQPFLLRNGANGFHEAVGEIWSLAPPRKILSGILSPDFQEDNEFEI 418
 DB 889 GHIOYDMAYAAQPFLLRNGANGFHEAVGEIWSLAPPRKILSGILSPDFQEDNEFEI 418
 QY 419 NFLKQALALTYGLPTVYLMKEMRMVVEGELPKQOMKMKMKREIYGVVEPYPHDET 478
 DB 948 NYLMSIALDKIAFLPFGYILMDQRMKVFQGRKIDEDYNOQNMNLRKLYGCLPVPYRSED 1007
 QY 479 YCDPASLFFHSNDYSFIRYRTTLYQFOFQALCOAKHKGEPHLCIDISNSTEAGOKLFNM 538
 DB 1008 DFPDGAFFHHPANVPYRFRVSIIOFQHEALCRAAGTGPHKCDIYQSEKAGLADA 1175
 QY 539 MLRLGSEPTWTLAENVGAKNNVRLNTEPELFTWL--KDONKSEFVGM-STDMSP 595
 DB 1068 AMRLGSEPTWTLAENVGAKNNVRLNTEPELFTWL--KDONKSEFVGM-STDMSP 595
 QY 596 ADOSIKYRISLKSALG-----DKAEWMDNEMTLFRSSVAYAMQRYFLAKY 641
 DB 1128 AVTEFHAAITDADFLGMSVGTQKATAGAW-----VLLALALVELLITISIFLYAK 1175
 DB

RESULT 7

JC2038
 peptidyl-dipeptidase A (EC 3.4.15.1) - rat
 N:Alternate names: angiotensin converting enzyme; kinase II
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: JC2038
 R:Author: Krieger, J.E.; Jacob, H.J.; Mukoyama, M.; Pratt, R.E.; Dreu, V.J.
 Biochem. Biophys. Res. Commun. 196, 380-386, 1994
 A:Title: Angiotensin converting enzyme and genetic hypertension: Cloning of rat cDNAs
 A:Reference number: JC2038; MUID:94121658; PMID:8292044
 A:Accession: JC2038
 A:Molecule type: mRNA
 A:Residues: 1-1313 <COI>
 A:Cross-references: GB:U03734; NID:9437289; PIDN:AAA8211.1; PID:9437290
 A:Note: the authors translated the codon AGC for residue 159 as Tyr
 C:Comment: This enzyme is a zinc-containing dipeptidyl carboxypeptidase that cleaves angiotensin
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; peptidyl-dipeptide hydrolase; transmembrane protein;
 F:393-400,990-998/Region: catalytic #status predicted
 F:1264-1284/Domain: transmembrane #status predicted <TM>

Query Match 33.9%; Score 1310; DB 1; Length 1313;
 Best Local Similarity 42.0%; Pred. No. 4.8e-84;
 Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps 7;

2 TIEBOAKTFLEKRNHEADLFYQSSLSAMNYNTNTEENYOMNNAAGKMAFLKEOSTL 61
 DB TIEBOAKTFLEKRNHEADLFYQSSLSAMNYNTNTEENYOMNNAAGKMAFLKEOSTL 61
 650 TDEAKANFVEEYEDTAKVLMNNEAEANMHNNTNITTEGSLILQKNKNSHTLKYGTW 709
 QY 62 AOMTPELOEIONLYVKLOALQONGSSVLSSEKSKRLNTLNTMSTYSTGKVCNPDPQ 121
 DB AKTEPVSNFQNSTIKRIKIKVONVRAVLPNNELEYNQILLDMETTYSVANVCYTG-- 767
 QY 710 AKTEPVSNFQNSTIKRIKIKVONVRAVLPNNELEYNQILLDMETTYSVANVCYTG-- 767
 DB 710 AKTEPVSNFQNSTIKRIKIKVONVRAVLPNNELEYNQILLDMETTYSVANVCYTG-- 767
 QY 122 ECLLEPGLNTEIANSLSLDYNERLWAMESREVEGKQLRPLEEYVVLKNEKARANHYEDY 181
 DB 768 TCLSLPEPLTNIMATSRKYEELLMWMSKMDKVRALIPFPKYVDVSNKIAKLNGYSDA 827
 QY 182 GDYMGDEYVNGVGYDYSRGQLIEDVHTFEIKRPLYEHLHAAYRAKIMAYPS-YISP 240
 DB 828 GDSMRSSESDLE-----QDEKLYQELQPLTLNHAAYRSLRHRTSEYINL 877
 QY 241 IGLPFAHLGDMGRFWNTLSLTPFGQKPNIDVTAMVDQAMDARIFEAKKFFVSV 300
 DB 878 DGPIDPAHLGDMGRFWNTLSLTPFGQKPNIDVTAMVDQAMDARIFEAKKFFVSV 300
 QY 301 GLPMTQGFENSMILTPRGVOKAVCHPTAMDLSKG-DFRILMCTKYVMODELTAHHMG 359
 DB 938 GLPMTQGFENSMILTPRGVOKAVCHPTAMDLSKG-DFRILMCTKYVMODELTAHHMG 359
 QY 360 HIQDYMAVAAPFLIRNGANGFHEAVGEIMSLSAATPKHLKSLGLSPQEDNEIN 419
 DB 998 HIQDYMAVAAPFLIRNGANGFHEAVGEIMSLSAATPKHLKSLGLSPQEDNEIN 419
 QY 998 HIQDYMAVAAPFLIRNGANGFHEAVGEIMSLSAATPKHLKSLGLSPQEDNEIN 419
 DB 420 FLKQALITVGLTPEYTMLEKRWAVKGEIIPKQDMKRWEMKREITGVVEPYPHET 479
 QY 1057 FLKQALITVGLTPEYTMLEKRWAVKGEIIPKQDMKRWEMKREITGVVEPYPHET 479
 DB 1117 FDEPSKTHVPRANVYIRYIFSLTIOFQFHEALCKAAGHTGPKCDIYQSKREAKLADA 1176
 QY 480 CDPASLFVSNVDSYFIRYTRTLQFQFOEALCOAAHBEPLHCDISNSTEAGKFLPM 539
 DB 1117 FDEPSKTHVPRANVYIRYIFSLTIOFQFHEALCKAAGHTGPKCDIYQSKREAKLADA 1176
 QY 540 LRLGKSEPTLALENVGAKNMNVRPLNTEPEPLTWMKQDN--NSFVGM-STDPMS 594
 DB 1177 MRLGSKSEPTLALENVGAKNMNVRPLNTEPEPLTWMKQDN--NSFVGM-STDPMS 594
 QY 1177 MRLGSKSEPTLALENVGAKNMNVRPLNTEPEPLTWMKQDN--NSFVGM-STDPMS 594
 DB 1177 MRLGSKSEPTLALENVGAKNMNVRPLNTEPEPLTWMKQDN--NSFVGM-STDPMS 594

RESULT 8
 A34402
 pepitdyl-dipeptidase A (EC 3.4.15.1) precursor, testicular - rabbit
 N:Alternate names: angiotensin I-converting enzyme; dipeptidyl carboxypeptidase I; pepit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text-change 10-Sep-1999
 R:Accession: A34402; A60724; A56232; C18700
 R:Kumar, R.S.; Kusari, J.; Roy, S.N.; Softer, R.L.; Sen, G.C.
 J. Biol. Chem. 264, 16754-16758, 1989
 A:Title: Structure of testicular angiotensin-converting enzyme. A segmental mosaic isozy
 A:Reference number: A34402; MUID:9380303; PMID:2550457
 A:Accession: A34402
 A:Molecule type: mRNA
 A:Residues: 1-737 <KMD>
 A:Cross-references: GB:J05041; NID:g164744; PIDN:AAA31153.1; PID:g164745
 R:Sen, G.C.; Thekkumkara, T.J.; Kumar, R.S.
 J. Cardiovasc. Pharmacol. 16(Suppl.4), S14-S18, 1990
 A:Title: Angiotensin-converting enzyme: structural relationship of the testicular and the
 A:Reference number: A60724; MUID:9155372; PMID:1705622
 A:Accession: A60724
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 73-173 <SEN>
 A:Note: Identical sequences were obtained for mRNAs from lung and testes
 R:Chen, Y.N.P.; Rjordan, J.F.
 Biochemistry 29, 10493-10498, 1990
 A:Title: Identification of essential tyrosine and lysine residues in angiotensin convert

A:Reference number: A36232; MUID:91104959; PMID:2176870
 A:Accession: A36232
 A:Molecule type: Protein
 A:Residues: 154-160:236-242 <CHE>
 R:Ikata, K.; Iai, C.Y.; El-Dorri, H.A.; Softer, R.L.
 Biochem. Biophys. Res. Commun. 107, 1097-1103, 1982
 A:Title: The NH₂- and COOH-terminal sequences of the angiotensin-converting enzyme iso
 A:Reference number: A90107; MUID:83048249; PMID:6291514
 A:Accession: C18700
 A:Molecule type: protein
 A:Residues: 33-35,'SN',38-39,'SS','FAEL',73 <TWA>
 A:Note: Several of the amino acids in reported are tentative
 C:Comment: The pulmonary and testicular isoforms of this enzyme differ substantially
 C:Superfamily: the two isoforms arise by alternative splicing of one gene.
 C:Keywords: alternative splicing; peptidyl dipeptide hydrolase; testis; transmembrane

Query Match 33.2%; Score 1283; DB 1; Length 737;
 Best Local Similarity 40.8%; Pred. No. 1.6e-82;
 Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;

2 TIEBOAKTFLEKRNHEADLFYQSSLSAMNYNTNTEENYOMNNAAGKMAFLKEOSTL 111
 DB TIEBOAKTFLEKRNHEADLFYQSSLSAMNYNTNTEENYOMNNAAGKMAFLKEOSTL 111
 75 TDEAKANFVEEYEDTAKVLMNNEAEANMHNNTNITTEGSLILQKNKNSHTLKYGTW 134
 QY 52 SAFLEKROSTIAOMYPIQEIOLNLYVKLOALQONGSSVLSSEKSKRLNTLNTMSTYST 111
 DB 135 SAFLEKROSTIAOMYPIQEIOLNLYVKLOALQONGSSVLSSEKSKRLNTLNTMSTYST 111
 QY 112 GKYCNPNPNQECILDEPLNTEIANSLSLDYNERLWAMESREVEGKQLRPLEEYVVLKNE 171
 DB 185 ANYCVRDYG--SCILDEPLNTEIANSLSLDYNERLWAMESREVEGKQLRPLEEYVVLKNE 171
 QY 172 MARANHEYDGYWBRDYEYNGVGYDYSRGQLIEDVHTFEIKRPLYEHLHAAYRAKIM 231
 DB 243 AALNNGYVADGDSKRSKMTETPTLE-----QDEKLYQELQPLTLNHAAYRSLRHRTSEYINL 292
 QY 232 MAY-PSYISPTGLPFAHLGDMGRFWNTLSLTPFGQKPNIDVTAMVDQAMDARIFE 290
 DB 293 RHYGAQHINLEGPFAHLGDMGRFWNTLSLTPFGQKPNIDVTAMVDQAMDARIFE 290
 QY 291 KEAKRFFVGVGLPMTQGFENSMILTPRGVOKAVCHPTAMDLSKG-DFRILMCTKYVM 349
 DB 353 KEAKRFFVGVGLPMTQGFENSMILTPRGVOKAVCHPTAMDLSKG-DFRILMCTKYVM 349
 QY 353 KEAKRFFVGVGLPMTQGFENSMILTPRGVOKAVCHPTAMDLSKG-DFRILMCTKYVM 349
 DB 353 KEAKRFFVGVGLPMTQGFENSMILTPRGVOKAVCHPTAMDLSKG-DFRILMCTKYVM 349
 QY 350 DEFLAHENGHIQDYMAVAAPFLIRNGANGFHEAVGEIMSLSAATPKHLKSLGLSPQ 409
 DB 413 DEFLAHENGHIQDYMAVAAPFLIRNGANGFHEAVGEIMSLSAATPKHLKSLGLSPQ 409
 QY 410 FQEDNETEINFLKQALITVGLTPEYTMLEKRWAVKGEIIPKQDMKRWEMKREITGV 469
 DB 473 FQEDNETEINFLKQALITVGLTPEYTMLEKRWAVKGEIIPKQDMKRWEMKREITGV 469
 QY 470 VEPVPRDEYVCDPASLFVSNVDSYFIRYTRTLQFQFOEALCOAAHBEPLHCDISNST 529
 DB 532 CDPASLFVSNVDSYFIRYTRTLQFQFOEALCOAAHBEPLHCDISNSTEAGKFLPM 539
 QY 530 TEAGKRLFMRLGKSEPTLALENVGAKNMNVRPLNTEPEPLTWMKQDN--NSFVGM 587
 DB 592 KEAGKRLADAMKLGSKPPEAMKYITGPNMSASAMNYFRPLDMDLITENGREGKILG 651
 QY 588 W-STDPMSPYADS 599
 DB 652 WQYTWTPNSARS 664

RESULT 9
 S35484
 pepitdyl-dipeptidase A (EC 3.4.15.1) precursor, pulmonary splice form - rabbit
 N:Alternate names: angiotensin-converting enzyme; dipeptidyl carboxypeptidase I; kin
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text-change 10-Sep-1999
 R:Accession: S35484; A23455; A18700; A38655; A49726; S17509

R:Thekkumkara, T.J.; Livingston III, W.; Kumar, R.S.; Sen, G.C.
 Nucleic Acids Res. 20, 683-687, 1992
 A:Title: Use of alternative polyadenylation sites for tissue-specific transcription of the
 A:Reference number: S35484; MUID:92178960; PMID:1311831
 A:Accession: S35484
 A:Molecule type: mRNA
 A:Residues: 1-1309 <THE>
 A:Cross-references: EMBL:X62551
 R:Itawa, K.; Blacher, R.; Soffer, R.L.; Lal, C.Y.
 Arch. Biochem. Biophys. 227, 188-201, 1983
 A:Reference number: A23455; MUID:84051289; PMID:6314908
 A:Accession: A23455
 A:Molecule type: protein
 A:Residues: 34-47, 'N', 49-55 <IMA>
 A:Experimental source: lung
 R:Itawa, K.; Lal, C.Y.; El-Dorri, H.A.; Soffer, R.L.
 Biochem. Biophys. Res. Commun. 107, 1097-1103, 1982
 A:Title: The NH₂- and COOH-terminal sequences of the angiotensin-converting enzyme isoform
 A:Reference number: A90107; MUID:83048249; PMID:6291514
 A:Accession: A18700
 A:Molecule type: protein
 A:Residues: 34-44; 754-755, 'L', 757 <IM2>
 R:Kumar, R.S.; Thekkumkara, T.J.; Sen, G.C.
 J. Biol. Chem. 266, 3854-3862, 1991
 A:Title: The mRNAs encoding the two angiotensin-converting isoforms are transcribed from
 A:Reference number: A38655; MUID:91139683; PMID:1847388
 A:Accession: A38655
 A:Molecule type: DNA
 A:Residues: 1-88 <KUM>
 A:Cross-references: GB:M58579
 R:Ramchandran, R.; Sen, G.C.; Misson, K.; Sen, I.
 J. Biol. Chem. 269, 2125-2130, 1994
 A:Title: Regulated cleavage-secretion of the membrane-bound angiotensin-converting enzyme
 A:Reference number: A49726; MUID:94124568; PMID:9294466
 A:Accession: A49726
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1236-1258 <RAM>
 A:Experimental source: testis
 R:Kirley, T.L.
 Biochem. J. 278, 375-380, 1991
 A:Title: The Mg(2+)-ATPase of rabbit skeletal-muscle transverse tubule is a highly glyco
 A:Reference number: S17509; MUID:91378880; PMID:1654880
 A:Accession: S17509
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 34-55 <KIR>
 C:Comment: This enzyme converts angiotensin I to angiotensin II in presence of divalent
 ver. the enzyme has been found also in renal tubules and intestinal mucosa.
 C:Superfamily: mammalian peptidyl-di-peptidase A
 C:Keywords: alternative splicing; blood pressure control; chloride; glycoprotein; intest
 F:1-33/Domain: signal sequence #status predicted <SIG>
 F:34-1309/Product: peptidyl-di-peptidase A, pulmonary #status experimental <MAT>
 F:59/79,150,222,448,512,680,698,717,945,1194/Binding site: carboxylate (Asn) (covalent)
 Query Match 33.2%; Score 1283; DB 1; Length 1309;
 Best Local Similarity 40.8%; Pred. No. 3.8e-82;
 Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;
 2 TIEQAKTFLDKFNHAEEDLFYQSSLASWNTNTITEE-----NVQNMNN--AGDKW 51
 Db 647 TDEAESRVEEEDRFOQVWMEYAMANNNTNTITEESKILLQKNQIAHHTLTYLQW 706
 52 SAELEQSTLAQMYPIQETQNTLVKIQALQALQOQSSVSEDKSKRLNTILTMSTIYST 111
 Db 707 -----ARRFDVSNFQNTSKRIKKYQDLQRAVLPVKELEBYNQILLDMETISV 756
 112 GKVCNDNDPQECLELPEGLMEIMANSNDYNERLWAMESRSVQKQRLPYEEYVAKKE 171
 Db 757 ANCCRVAG--SCIQLEPDLTNLMATSRKDELIMWTSVRDYGRIILPEFYVEFTTK 814
 172 MARANHYEDYGDWGRDYEYNGVDGYDSRGQILIEDVETFEIKPLYEHLTAAYRAKIM 231

Db 815 AARLNGYVADAGDSWRMYETPTLE-----QDLERLEQLOPLYINLHAAYGRAH 864
 Oy 232 NAY-PSYISPGICPAHLLGDMGRFTNLSLTFEGQKPNIDVDAMVDQAQRTF 290
 Db 865 RHVQAHINLEGPPIALLGNMNAQWTSNIYDLVAPPPASSTMDATAMIKQWTRRF 924
 Oy 291 KEAEKFFVSVGLPMTQGEWNSMLDPCVQAVCPAPWADLKG-DERIIMCTRYMD 349
 Db 925 EADKFFISGLLEVPPEFNKSKMLEKPTGREVYCHASAMDPYNGKDFIKCTTYNME 984
 Oy 350 DELFAHHEMHIOYDAMVAAOPFLRNGAEGHEHVEGEMTSATPHTKISGLISPD 409
 Db 985 DLVVHHEMGHIOFYQYKQDLPALRGAPNGHEALGDLALSVSPHLHSINLSSS 1044
 Oy 410 FOEDNEFEINFLKQALTIYGLPTTYMLKRWMEYFGEIPKDDMKKWEKREIVG 469
 Db 1045 -GGGEHNDINPMALDKIAIFIPSYLVDEMRRVDSITKENYNOEWMSRLAKYQGL 1103
 Oy 470 VEPYPHDETICDPASLSEHVSNDSFTYRTYRITLQPOFQALCOAKHSGPLKDISNS 529
 Db 1104 CPAPRSGQDFDPGAKFHPSVYIRYFVSFTIQFQFHALCKAGHGTPLTCDIYOS 1163
 Oy 530 TEAGCKLFNMLRLGKSEPTLALENVYGAKNMVRPLNYPEPLTWLKDQ--KNSFVG 587
 Db 1164 KEAKRLADAMKLGISKPWEBAKVITGQFNMSASAMNYFKPLMDMLTENGREGKIG 1223
 Oy 588 W-STDWSPTADQS 599
 Db 1224 WQYTWTPNSARS 1236
 RESULT 10
 S65472
 peptidyl-di-peptidase A (EC 3.4.15.1) precursor - horn fly
 N:Alternate names: angiotensin I-converting enzyme
 C:Species: Haematobia irritans (horn fly)
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 22-Jun-1999
 C:Accession: S65472; S65431
 R:Willemsen, G.; FitzGerald, C.; Gough, J.; Ridding, G.; Elvin, C.; Kemp, D.; Willemsen
 Eur. J. Biochem. 237, 414-423, 1996
 A:Title: Cloning and characterisation of angiotensin-converting enzyme from the dipe
 A:Reference number: S65431; MUID:96215437; PMID:8647080
 A:Accession: S65472
 A:Molecule type: DNA
 A:Residues: 1-611 <WID>
 A:Cross-references: EMBL:L43965; NID:9908759; PID:AAA70427.1; PID:9908760
 A:Note: the source is designated as Haematobia irritans exigu
 A:Accession: S65431
 A:Molecule type: protein
 A:Residues: 18, 'P', 20-42; 75-77, 'Q', 79-81, 'X', 83-84; 179-190 <WIM>
 A:Note: the source is designated as Haematobia irritans exigu
 C:Gene: ACE
 C:Superfamily: mammalian peptidyl-di-peptidase A
 C:Keywords: glycoprotein; metal binding; peptidyl-di-peptide hydrolase; zinc
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-611/Product: peptidyl-di-peptidase A #status predicted <MAT>
 F:53,196,531/Binding site: carboxylate (Asn) (covalent) #status predicted
 Query Match 27.3%; Score 1054.5; DB 2; Length 611;
 Best Local Similarity 37.7%; Pred. No. 1.6e-66;
 Matches 222; Conservative 105; Mismatches 229; Indels 23; Gaps 8;
 1 STIEQAKT-FLDKFNHAEEDLFYQSSLASWNTNTITEE-----NVQNMNN--AGDKW 59
 Db 18 AFKEIVATETVLQINNELAKHTNVEEVSAAVSNNTDNERLRNLSAENAKFLKEVA 77
 Oy 60 TLAQMYPIQETQNTLVKIQALQALQOQSSVSEDKSKRLNTILTMSTIYSTVCPDN 119
 Db 78 KDIOQFNKRTGYSADVROQFKSLKSTGYSLAPADYELLLEVISAMESNPAKYRVCDYK 137
 Oy 120 PQEC-LLEPGLNEMANSNDYNERLWAMESRSVQKQRLPYEEYVAKKEANAHY 178

Db 138 SACCDLSLDEIEIITKSDPEELKYWTQFYDKAGTPRNSFEKVELNITSKALINNE 197
 QY 179 EDYGDYRGDYEVNGVDYDSRGQIED-VENHFEIKPLYEHLAAVRAKIMANAY-DS 236
 Db 198 TDGAELWLDYE-----DAFEEQLEAIFEDIKPLQVIGYARILNKRYGDE 246
 QY 237 YISPIGCLPAILGDMKRGWTLULSYLVEFGKPNIDVTADAMVDAQIRFEAKERF 296
 Db 247 VVSKTGPLPMLHGLGNMAAQMSIADIVSPFEKPLVDYSDKVAAGTIPRLKMFQGDGF 306
 QY 297 FVVGSLPNTMTOGFENSMILDPGNVOKAVCHPTAMUG-KGDFRILMKRYMDDPLTAH 355
 Db 307 FQSMGKLKLPQEWKWSILEKPDGDLCHASAMFPYITLDDYRIKQCTRYVQDQFPTVH 366
 QY 356 HENGHIQYDMAVAAPPELLRNGANGFHEAVGEIMSLAATPKHLKISGLSPDQEDNE 415
 Db 367 HENGHIQYFLOYOHQFVYRTGANPGFHEAGVDLSLSTPKHLERVELIK-NYSDNE 425
 QY 416 TEINPELLKQALITVGLPTPLMLEKRRNVEFGKELPKDQMKMKWEMKREIVGVEPVH 475
 Db 426 ARINQLFLTALDKIVLPFAFTMDKRYMALFRQADKSMNCAFWRILREYSGIEPPVVR 485
 QY 476 DETYCDPASLFEVNSDYSEIRYTRTYLQFOFQALCOAA-----KHGGLHKODISNS 529
 Db 486 TEKDFAPAKYHVSADVEYLRIVYSITIQFQYKSCACTAGERYVNOETPLDNCIDYGS 545
 QY 530 TEAGQKLFNMLRGKSEPTLALENVGAKNNVPRLLNYFEPRLTWLK 578
 Db 546 KEAGKLFENMLSLGASKPMPDALLAENGERMTMGKALHAYFELRWLE 594

RESULT 11

angiotensin-converting enzyme-related protein - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:date: 28-May-1997 #sequence_revision 18-Jul-1997 #text_change 21-Jul-2000
 C:Accession: J05374
 R:Taylor, C.A.M.; Coates, D.; Shlirsa, A.D.
 Gene 181, 191-197, 1996
 A:title: The Acer gene of Drosophila codes for an angiotensin-converting enzyme homolog
 A:reference number: J05374; MUID:97128790; PMID:8973330
 A:Accession: J05374
 A:Molecule type: mRNA
 A:Residues: 1-630 <TAY>
 A:Cross-references: EMBL:X96913; NID:g1405881; PIDN:CA65632.1; PID:g1405882
 C:Genetics:
 A:Gene: Acer
 C:Superfamily: mammalian peptidyl-di-peptidase A

Query Match 26.6%; Score 1027.5; DB 2; Length 630;
 Best Local Similarity 36.0%; Pred. No. 1.4e-64;
 Matches 215; Conservative 110; Mismatches 250; Indels 23; Gaps 9;
 QY 6 QAKTFLDKNHEADLFYQSSLASWNTNTITEENVQNMNAGDWSAFLKQSTLAQ 65
 Db 33 EARREFELENEQLRFRHEEFLSGVNTYNTVEARQAMIEVYANALNKLRAQITSS 92
 QY 66 PLQELQNTLVKIQALQALQONGSSVSEDSKSKRLNTILMTGSTYSGVCPNDPQC-L 124
 Db 93 DYVQSEDDIRQOAEHLKSLGASALNADDYLAQNAISSMOTNTATATVCSYTNRSCSL 152
 QY 125 LLEPGLNETMANSIDYNERLWAMESRSEVGKQLRPLVEEYVVLKNEARAHNYEDY 184
 Db 133 TLEPHIGRLSHSRPALAWYRWMHDKSGTPMKNQFAEYVRLKRSQSLGHRSTADY 212
 QY 185 WRGDEVNGVDYDSRGQLLEDVHTTEIKPLYEHLAAVRAKIMANAY-PSYISPIGC 243
 Db 213 WVOFYE-----DPPDER-----OLDATPEKQLPLRYRQLHGVYRFLRQHGYPDVPAGSN 262
 QY 244 LPAHLIGDMGWFNTNLSLVVPGQKNIDVTADAMVDAQIRFEAKERFVYSGLP 303
 Db 263 IPISLILGNMGOSNELLDTFTPYEPKFPVYKAMEKQGYVOKLFEIGDOFQFOSLGR 322

QY 304 NMTQGFENSMILDPGNVOKAVCHPTANDLQK-GDFRILMKRYMDDPLTAHHEMHIQ 362
 Db 323 ALPSPFWNLSTVLRPDD-RQVYCHASAWDFYQDSVRIKMCLEVDSHFYVYVHHLGHIQ 381
 QY 363 YDMAVAAPPELLRNGANGFHEAVGEIMSLAATPKHLKISGLSPDQEDNEETINPL 422
 Db 382 YLYQEQGPAYVRGAPNPFGEHAGVDYALASMSKHLKALIGLE-NORDEKSRINDLE 440
 QY 423 KQALITVGLPTPLMLEKRRNVEFGKELPKDQMKMKWEMKREIVGVEPVHDEYCDP 482
 Db 441 KQALSKIVYLPFGYAVDKRYAVFERNELDESQWNCQFQMSRSEGVGPVFRTEKDRDP 500
 QY 483 ASLFVNSDYSEIRYTRTYLQFOFQALCOAAKHEP-----LHKODISNTEAOKL 536
 Db 501 PAKYIHDADVETLRFAAHIFQFQHKYLCRAAGYAPNNSRLTLDNCIDIGSKAAGSL 560
 QY 537 FNMRLRGKSEPTLALENVGAKNNVPRLLNYFEPRLTWLKDQNKSFVSWTDSWP 594
 Db 561 SQFLSKGNSRHKVEYLEEFTGETEMDPALLLEYFELVOWLKE-NSRLGVPLGMGP 616

RESULT 12

peptidyl-di-peptidase A (EC 3.4.15.1) 67k precursor - fruit fly (Drosophila melanogaster)
 N:Alternate names: angiotensin-converting enzyme
 C:Species: Drosophila melanogaster
 C:date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 26-Feb-1998
 C:Accession: A57533
 R:Corneil, M.J.; Williams, T.A.; Lamango, N.S.; Coates, D.; Corvol, P.; Soubrier, F.;
 J. Biol. Chem. 270, 13613-13619, 1995
 A:title: Cloning and expression of an evolutionary conserved single-domain angiotensin
 A:reference number: A57533; MUID:95293950; PMID:775412
 A:Accession: A57533
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-615 <COR>
 A:Cross-references: G8:U25344
 A:Gene: Flybase:Ance
 A:Cross-references: Flybase:FBgn0012037
 C:Superfamily: mammalian peptidyl-di-peptidase A
 C:Keywords: peptidyl-di-peptide hydrolase

Query Match 26.4%; Score 1022; DB 2; Length 615;
 Best Local Similarity 35.6%; Pred. No. 3.3e-64;
 Matches 213; Conservative 118; Mismatches 245; Indels 22; Gaps 9;
 QY 4 EEOAKTFLDKNHEADLFYQSSLASWNTNTITEENVQNMNAGDWSAFLKQSTLAQ 63
 Db 22 EIOAKLEYLENLKEIARNTVETEAWAAGSNTIDENKKNEISAEIAKFEVASDTE 81
 QY 64 MYPLQELQNTLVKIQALQALQONGSSVSEDSKSKRLNTILMTGSTYSGVCPNDPQC 123
 Db 82 KFORKSTVSEELKQFAIKLKYALRPEDDYALLDTLSLAMESNEFKVAVCYKQSTK 141
 QY 124 LLEPGLNETMANSIDYNERLWAMESRSEVGKQLRPLVEEYVVLKNEARAHNYEDY 182
 Db 142 DLALDPELEIEYISRSDBELAYWRFYQKAGAVASQDERVEYELNTAKAKLNFTSCA 201
 QY 183 DYMGQDEVNGVDYDSRGQLLEDVHTTEIKPLYEHLAAVRAKIMANAY-PSYISPT 241
 Db 202 EAWLDYE-----DFFEQQLEDI--PADIRPLLPADPWLCAFRILRKHGYDAVAVSET 251
 QY 242 GCLPFAHLIGDMGWFNTNLSLVVPGQKNIDVTADAMVDAQIRFEAKERFVYSG 301
 Db 252 GPIPMHLLGNMGWQMSSEINDIYSPFEKPLVDYSAMEKQATPLPKAFMGDDFTTSN 311
 QY 302 LPAHTQGFENSMILDPGNVOKAVCHPTANDLQK-GDFRILMKRYMDDPLTAHHEM 360
 Db 312 LTKLPLDQFPMHLLGKSTIEKPTDRDLVCHASAWDFYLDIVRIKQCTRYVQDQFVHHEIGH 371
 QY 361 IOYDMAVAAPPELLRNGANGFHEAVGEIMSLAATPKHLKISGLSPDQEDNEETIN 420

Db 372 IOEFLQYQHGFYRTGANGFHEANGVLSVSTPKHLEKIGLILK-DYVEDDEARINQ 430
QY 421 LKQALITVGTLPFTYMLEKRWMTFKGIPDQMKKWKREIVGVVEVPDEHYC 480
Db 431 LELTALDKLTVLPEAFMDKRYMSLFRGEVDKANNKCAFWKLDEYSGTEPPVNSKDF 490
QY 481 DPASLPHVNDSEFIYRTYLYOFQFOALC-QAKH-----EGPLKCDISNSTEQ 534
Db 491 DAPAKHISADVEYELKLVSLIFQOFYKACIKAGQYDPDVELPLDQCDIYSARAGA 550
QY 535 KLEFNLRLCKSEFWTALENVGAKNNVPLNTPEPLTYLQKON-KNSFWGWT 590
Db 551 AFHNLMSGSKRPDLEAFNGERIMSGKATAEFEPLRVMLAEENIKNNVHIGWT 608

RESULT 13

T15792

hypothetical protein C42D8.5 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000

C:Accession: T15792

R:Hallsworth, K.

submitted to the EMBL Data Library, April 1996

A:Description: The sequence of *C. elegans* cosmid C42D8.

A:Reference number: Z18405

A:Accession: T15792

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1907 <HAL>

A:Cross-references: EMBL:056966; NID:g1293844; PID:g1293847; PIDN:AAA98719.1; GSPDB:GN00

A:Genetic source: strain Bristol N2; clone C42D8

A:Gene: CESP:C42D8.5

A:Map position: X

A:Insertions: 140/3; 194/3; 300/2; 467/3; 551/2; 600/2; 697/3; 774/2; 851/3

Query Match

Best Local Similarity 27.0%; Score 638.5; DB 2; Length 907;

Matches 167; Conservative 132; Mismatches 280; Indels 39; Gaps 18;

QY 4 EEOAKTFIDKFNHEADLFYQSSLASWNTNTNTEENYQNMNAGKWSAFLEQSTLQ 63
Db 178 EEKLRSLAGYEAEAIKLVREVALSGWRIFNDASPSIKALDEAEVLTTFVASTMQAK 237
QY 64 MYPLQEIQLTVKLOLQALQONGSSVLSSEDSKRLTILNTSTIYSTGKVCNPNPQEC 123
Db 238 QPDMASVTDEKVRKOLGYVSEGMASALAPSRFADYSQAQALNRDSDSTICDKVPPPC 297
QY 124 LLEPGINETMANSLDNERIMAWESKRSYVQOLRLRYEYVVLANKEMARAHYDYGD 183
Db 298 ALQKIDMDSIFRRKAKASRLQHLWVSVTAIKRS-KPSYNNITITISNEGKILNGFANGGA 356
QY 184 YMRDYEANG-VQGYDYSRGQLEDEYHTEFEIKPLYEHLHAYVRAKIMNAV--PSYISF 240
Db 357 MMSAFLPMSKVKHAEF---DLNKQIDKISTYQPFQOLHATMRQGLAGITSNPGLSK 413
QY 241 IGCPLPAHLGDMGRTNLYSLTVPGQKPNIDVTDAVD---QAMDAQIFKEAEK 296
Db 414 DGPIPAHLFGSLDGGMSAHYEQTKEPEEES--ETPEAMLSAFNTQYITTKKMFVAYRY 471
QY 297 FVSVGLPNNMNGEFNEMSLDPPGVQKAVCHP-TAMDL-GKGPRLIMCKYVMDDPLTA 354
Db 472 FKSGFPHLPKSYTSSIFARWVS-KMICHPAALADWRAPNDRFVAAACQQLDEPDEQA 530
QY 355 HHEGHILOYDAYAAOFPLLRNGANEGFHEAVGEIMSLSATPKHLKISGLSPDEQDN 414
Db 531 HSLVQYTYQYLYLDQSLFRQASPVITDAIANAFHLSSTNPHYLLSQGLVSEHLDIK 590
QY 415 ETE-INFLKQALITVGLDPTTYLKEKRWMAVEKEIIPKQMKKWKREIVGVVEPV 473
Db 591 DSVIINKLYKESLESFETLFTIADMRWRELEDEGTVPKKNLNDRWMEINKKEGVASPO 650
QY 474 PHDETQCDPASLFR--VSNDYSFIRYTRTL---YQFOFQALCQAA---KHGFLMK 524

Db 651 PYNTSND--ALHNSVSQVHS---PATRLLISVYLFOILKALCORELFWLSEG----C 701
QY 525 DISNSTAGOKLFNMLKSEPTWLTALENVGAKNNVPLNTPEPLTYLQKON--582
Db 702 ILSEDTT--EKLRRTKMGSSITWKLALEMI SGKGEIDAPLLEYEPLNMLRNINED 759
QY 583 NSFVGSWTDSPPYADOST 600
Db 760 QVYVGMGEGEPTVEEI 777

RESULT 14

C83696

hypothetical protein BH0371 [Imported] - *Bacillus halodurans* (strain C-125)C:Species: *Bacillus halodurans*

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: C83696

R:Yakami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: C83696

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-532 <STO>

A:Cross-references: GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BAB04090.1; GSPDB:G

A:Experimental source: strain C-125

A:Gene: BH0371

Query Match

Best Local Similarity 4.1%; Score 157; DB 2; Length 532;

Matches 118; Conservative 83; Mismatches 213; Indels 144; Gaps 29;

QY 4 EEOAKTFIDKFNHEADLFYQSSLASWNTNTNTEENYQNMNAGKWSAFLEQSTLQ 59
Db 3 EODIRFLSEQKRVEDLYQVYLNHNVVATTEGQSDSKHQSLEYMAHFSRSEFOK 62
QY 60 -----TLQMTPLQEIQLTVKLOLQALQONGSSVLSSEDSKRLTILNTSTIY 109
Db 63 VTRFKRISLPLMQRRQDLHDHMKIKNQFE--EGTRQIQLSLE--KKISHVFTFPQOV 118
QY 110 SNGKVCNPNPQECLEPGLNETMANSLDNERIMAWESKRSYVQOLRLRYEYVVLK 169
Db 119 NSGRVSNNE-----LIDLIRYDLDEHRRQAWFA-SKEVGKTEKOLLQILRKR 166
QY 170 NEMARANYEDYGDYRGDYEVNGVQDYDSRGQLEDEYHTEFEIKPLYEHLNAV 226
Db 167 NEVARNLGEFT-----YHMSFAQGEIDLEQTFPMFETIKSSDQAFPMI 211
QY 227 -----RAKLMNAVSTYISPIGCLPAHLGDMGRTNLYSLTVPGQK-PNIDVTDA 278
Db 212 KDEIDEEKAVLKIKKDLRP-----WDVYDPEFQAPASIEHYD- 250
QY 279 MYDQAMDQRIKFEKAEKFVSGLPNNMNGEFNEMSLDPPGVQK-AVCHPTAMDLGKD 337
Db 251 -FDSFTKDDLEQVVSQFOAMELP--IDDLIKSDDLPRKKNKPFQFC--TDMD-RRGD 304
QY 338 FRLMCTKVYTMDFLTAHHEMGH-IQYDMAYAAOFPLLRNGANEGFHEAVGEIMSLSAT 396
Db 305 IYVLLNLQDSMTVMYVALLHEFGHAYYFKFIDRLPFLR-----FH-----SHLT 351
QY 397 PKHLKISGLS--PPOF-----DNET-----ELNFLKQALITVGLDPTTYLKEKRW 443
Db 352 EASALFFGRTKMAWYERFLGIDRETERIGRNEMKMLQROM-VVST-----RW 400
QY 444 NY-----FKG---EIPKQMKKWKREIVGVVEPVPHDETQCDPASLFRVVS-----N 490
Db 401 KLFESFEKSLYEDPDODINALMKLVKEIQYMAP--PEDIGSPMAKMHFSLAPYQY 458
QY 491 DY-----SFTRYTTRT 501

Db 459 DYLLGMAASQLHHYIKT 476

RESULT 15

AF1310

probable thermostable carboxypeptidases homolog lmo186 [imported] - *Listeria monocytogenes*C:Species: *Listeria monocytogenes*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AF1310

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maltournam, A.; Ma

ok, C.; Schluter, T.; Simoes, N.; Tiller, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AF1310

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-502 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAC99964.1; PTD:916411339; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo186

C:Superfamily: Thermus aquaticus carboxypeptidase Taq

Query Match

Best Local Similarity 4.0%; Score 154; DB 2; Length 502;

Matches 128; Conservative 101; Mismatches 215; Indels 192; Gaps 35;

```
2 TTEOAKTFIDKFNHEADLFYQSSLASWNT-----NITEENYONNNA 47
4 TLEEFYAIYIKKMALEAL-----ALVYMDLRTGAPAKMGESRDVIGVSEIEFNNQTS 59
48 GDKMSAPLKQSTLAQMPLOEIONLYKLOALQONGSSVLSDESKRL-----N 99
60 -EEMAFIAGLN-----QDENIS-EITKTLIE--SQKTYDLNKKIPKREAYEYK 106
100 TLTMTSTYTGKVCNPDNPOECLLEPGLNEIMANSIDYNERLMANESWSEYKQRL 159
107 KLVAAQETAMTARBN-----DFAFEPLIKIL-----136
160 PLYEEVVLKEMARAHYEDYGVWGRGVYNGVGY--DYSRQLIEDVHEFEIKP 217
137 -----EMK-----KFEYV--GYEENKIDTLIDYEPGVTVSLDSVEKVR- 177
218 LYEHLAAVYRAKLMNAVPSYISPIGCLPAHLIGDMWGRFTWLYSLVYFGOKPNIIDVT 277
178 --DGIMA--IREKIENE-----GVRPGDTILN 200
278 AMVDAMDAQRIFKEAEKFEVSVGLPNMTQGFWENSMLTDPGNVOKAVCHPTAMD LGKD 337
201 TKISRA-----KOKESIRI-LNKKGFDF-----EAGRLETV-HPEATGLNTGD 243
338 FRILMCTVMTDELTA-----HHEMGIHY-----DMAVAAPFLLRNANGSEFHEA- 387
244 VRI--TTRYNENDFKAAVFGTHIEGHAIYEONPAALVGP--LANGASMGIHESQSLF 299
368 -EIM--SLSAATPKHLKSGILSPDQ-----ED-----NETEINFLKQALTYVGLP- 433
300 YEIIIGSSIAFWKSNYADFOAITKPAFDQVKLEDFYRAVNISESSLRIEADTL--TYPL 357
434 ---FTYMLEKRWVWFKEIPRDOMKMKWMEKREIVGVPPVPHDEYCDPASLFHVN 490
358 HIMIRYELK--ALINGELEVKDLPKAWGDKYEYELGI--RDNDINGVLODIHNAAG 411
491 DYSFIRYTRTL-YOFQOALCOAAKHEGLHKCDISNSTEAGOKLF--NMLRLGK 546
412 DEGFYPSYALGLMTAAQFFNQM--QKEINIDAILASDYSSEKLTWLTGKHTFGK 467
547 PWTALALENVGAKNNVAPLINTFEPDLTWLKDQNK 582
```

Db 468 KPEILIDPTG-EGINPTYLIDLEKRYAYVQFNK 502

RESULT 16

AE1682

probable thermostable carboxypeptidases homolog lin199 [imported] - *Listeria innocua*C:Species: *Listeria innocua*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AE1682

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maltournam, A.; Ma

ok, C.; Schluter, T.; Simoes, N.; Tiller, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AE1682

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-502 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97229.1; PTD:916414500; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: lin199

C:Superfamily: Thermus aquaticus carboxypeptidase Taq

Query Match

Best Local Similarity 3.8%; Score 147; DB 2; Length 502;

Matches 124; Conservative 95; Mismatches 223; Indels 196; Gaps 33;

```
2 TTEOAKTFIDKFNHEADLFYQSSLASWNT-----NITEENYONNNA 47
4 TLEEFYAIYIKKMALEAL-----ALVYMDLRTGAPAKMGESRDVIGVSEIEFNNQTS 59
48 GDKMSAPLKQSTLAQMPLOEIONLYKLOALQONGSSVLSDESKRL-----N 99
60 -EEMAFIAGLN-----QDENIS-EITKTLIE--SQKTYDLNKKIPKREAYEYK 109
102 LNTMTSTYTGKVCNPDNPOECLLEPGLNEIMANSIDYNERLMANESWSEYKQRL 161
110 AQA-ETAMTARBN-----DFAFEPLIKIL-----136
162 YEEVVLKEMARAHYEDYGVWGRGVYNGVGY--DYSRQLIEDVHEFEIKP 219
137 -----EMK-----KFEYV--GYEENKIDTLIDYEPGVTVSLDSVEKVR- 177
220 EHLAAVYRAKLMNAVPSYISPIGCLPAHLIGDMWGRFTWLYSLVYFGOKPNIIDVTAM 279
178 DGIMA--IREKIENE-----GVRPGDTILN 202
280 VDAQMDAQRIFKEAEKFEVSVGLPNMTQGFWENSMLTDPGNVOKAVCHPTAMD LGDPR 339
203 ISEA-----KOKESIRI-LNKKGFDF-----EAGRLETV-HPEATGLNTGDVR 245
340 ILMCTVMTDELTA-----HHEMGIHY-----DMAVAAPFLLRNANGSEFHEA- 385
246 I--TTRYNENDFKAAVFGTHIEGHAIYEONPAALVGP--LANGASMGIHESQSLF 301
366 -----VGEIMSLAATPKHLKSGILSPDQ-----ED-----NETEINFLKQALTYVGLP 432
302 IIGSSIAFWKSNYADFOAITKPAFDVH-----LEDYRAVNISESSLRIEADTL--TY 355
433 P---FTYMLEKRWVWFKEIPRDOMKMKWMEKREIVGVPPVPHDEYCDPASLFHVN 488
356 PLHIMIRYELK--ALINGELEVKDLPKAWGDKYEYELGI--RDNDINGVLODIHNA 409
489 SNDSFIRYTRTL-YOFQOALCOAAKHEGLHKCDISNSTEAGOKLF--NMLRLGK 544
410 GDFYPSYALGLMTAAQFFNQM--QKEINIDAILASDYSSEKLTWLTGKHTFGK 465
545 SEPWTALALENVGAKNNVAPLINTFEPDLTWLKDQNK 582
```


Db 466 TKKPLELIDDTG-EGLNPTYLLELLEKRYAVYQFNK 502

RESULT 17
A12011
peptide synthetase [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: A12011
R:Kanezo, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, R.; Kanezo, T.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MID:21595285; PMID:11759840

A:Accession: A12011
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-987 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAW78013.1; PID:g17135467; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all1647

Query Match 3.6%; Score 139.5; DB 2; Length 987;
Best Local Similarity 19.6%; Pred. No. 0.15;
Matches 117; Conservative 80; Mismatches 208; Indels 191; Gaps 26;

QY 34 TTTTENVQNNMNAAGKSAFLKEOSTLAQMYPLQIYVLOALQOQSSVLSLD 93
DB 106 TLTTEHQVSOOLVACVQEAALPLQTVFLEGEPLGELISLT-----QITAEWQS-LSDD 159
QY 94 KSKRLTILMTSTIYSGKVCNPDN-----POECLLEPLGLNEMANSID 139
DB 160 ALELCNSPDDMLATITTSSTGRKGMNLNHRGYNRLTMQNTFSLQGDRAQRTSIC 219
QY 140 YNERLWAMESWSEVQKOLRPLVEEYVVLKEMARANHEDYGDYRGDEYVNGDGYIY 199
DB 220 FDISVWE-IFWTLMSGATTCIPQREVLVNFMEFAR-----WIOETQIVMHFVS 268
QY 200 SROGLIEDVHE-----TEFEIKPLEYEHHAAYRAKLMNAPSYISPICSLPAHLGDMWGR 256
DB 269 LFGFISALNETWSEFQKWLWLFSGRA-----LPHSFI-----QR 304
QY 257 W-----TNYSLTFVFGQKPNIDVTDAVDAQMAQ-----RIFEAKFFVS- 300
DB 305 WIDRHLGKTGLANLYGPT-----EASIDVYCHLITERPDERLTQIPIDGRAIDNYYVKYL 359
QY 301 --GLPNTQG-----FWNSM-----LTDCGNVOKAVCHPTAMDIGKD--FRILMCTVY 346
DB 360 DGGMQPQPGNMGELWLGVOALAGYLKDEKTRQAFCNPFTDI--PGDIYVTTGDLVE 418
QY 347 TMDDELTAHHEMGIHQDMAVYAAQPFLLRNGANGEGHEAVGEIIMSATPKHLSIGLL 406
DB 419 LPQSTIYH--GRIDHQAIV-----RGFIELGELIESVLTTPR--DYREAAAL 462
QY 407 SPDFQEDNETEINFLKQALTYGTLEPTTMLEKRWMTFKGELPKDQMKKWKWEKREI 466
DB 463 AVDGEQOKRLV-----ACISGKIKIKRFEKLEIQR-----494
QY 467 VGVVEPRPHDETYCDPASLFHVSNDYSFIYRTTLQFOFOEALQQAANHEBPLHKDI 526
DB 495 -----LPH--TWIP-----ORFLMLDSL--PKHNNGKLDKRAL 523
QY 527 -----SNSTAGQKLFNMLRIGKSEPTLALLENVVGAKNNVBPILLNYEPLETLW 577
DB 524 VTQLTSDSPSSPSSLPLPLPLGPAQRK-----LVVYFPPYQWL 562

RESULT 18
B82938
zinc metalloproteinase oligoendopeptidase F UU065 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000

C:Accession: B82938
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a
A:Reference number: A82870
A:Accession: B82938

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-608 <GLA>

A:Cross-references: GB:BA002106; GB:AF222894; MID:96899011; PIDN:AAF30470.1; GSPDB:GN
A:Experimental source: serovar 3; D10var 1
C:Genetics:
A:Gene: pepF-1; UU065
A:Genetic code: SGC3

Query Match 3.6%; Score 139; DB 2; Length 608;
Best Local Similarity 19.1%; Pred. No. 0.079;
Matches 132; Conservative 93; Mismatches 226; Indels 240; Gaps 32;

QY 4 EEOAK-----TFDDKFNHEAE-----DLFYQSSIASMYNTNTEENVQNMN 45
DB 28 EKKAATIAFPFPLDSKQNFQAMQILBEFTIVANRF--NYVSNMLNTNVYD-----78
QY 46 NAGDWSAFLKEOSTLAQMYPLQ-----EIQNLTVKLOALQOQSSV- 89
DB 79 --PKWNSW--SOKLSASFEELETAISNVNSVLANEAKIKIYELDSOLNVTYTKRYEIF 133
QY 90 -----LSEDKSKRLNTLNT--MSTYST-----KVCNPDNPOCLL 125
DB 134 KYQPTLLNNEOSKLESTILRADEGSTIFSTYTNNDKFSADAKGKPKIRKEALAFV 193
QY 126 LEPGLNEMANSIDYNERLWAMESWSEVQKOLRPLVEEYVVLKEMARANHEDYGDY 185
DB 194 HLKSDRYLRSKA-YLSMYRAYYDSRESITKM--LYYNYLSL--NQAKAKNFDDY--IA 246
QY 186 KQDYEVNGVDYDYSRGQIEDVEHTFEIKPLEYEHHAAYRAKLMNAPSYISPICLP 245
DB 247 KRAFP-----DVAQKSLITLIDYQVLYKDTNEDKRYANLYKLIV--SKIP-----295
QY 246 AHLGDMWGRFNTNLSLTFVFGQKPNIDVTDAVDAQMAQRTFEAKFFVSGLPNM 305
DB 296 -----WDNGLPL--ISKRIDPIBEAK-QMADLSLSLIGDE--YVS-----NI 333
QY 306 TQGFENSMKLDIDPGVQAACHPTAMDIGKGFRLIMCTKYTMDDELTAHHEMGIHQDM 365
DB 334 KRAFPDEKQVSWLPQKRGAGAYSIGTGKISKYYILMNTNSLRDYOYIVHELGHSMHS- 392
QY 366 AVAADPFLLRNGANGFHE--AVGEIMSLSAATPKHLSIGLLSPDFQEDNETEINFL- 421
DB 393 -----LYSNRTQKITSDYKIFYAELASIS-----NEYLYNYL 425
QY 422 -----LKQALTYGTLEPTTMLEKRWMTFK-----GEIPDQMKKWE 461
DB 426 LEKYKNDLEMKMILDEMISGFPAVTRQVIFSNFEWIANELINGSGAFADVWKEVOK 485
QY 462 MKREIVG--VYEP-----YPHDETYCDPASLFHVSNDYSFIYRTTLQFO 505
DB 486 LELETYNRPIVEDLNSIYSLSSITPLRILP-----FYGVNFY-----YKY 526
QY 506 QFOEALQQAANHEBPLHKDINSSTAGQKLFNMLRIG-----KSEPW 548
DB 527 AVGVAAAIISGR-----VFTKYVGAKQKAYDFELSSGSGKDPDITIKLLGVDLTKPQAW 580
QY 549 TALENVVGAKNNVBPILLNYEPLETLW 579
DB 581 OEALTEIV-----LWIKD 593

RESULT 19
D82881
zinc metalloproteinase oligoendopeptidase F UU521 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C/Accession: D82881
 R/Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
 submitted to Genbank, February 2000
 A/Description: The complete sequence of *Ureaplasma urealyticum*. Alternate views of a mit
 A/Reference number: A82870
 A/Accession: D82881
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-611 <GLA>
 A/Cross-references: GB:AE002150; GB:AF222894; NID:g6899515; PIDN:AAF0934.1; GSPDB:GN001
 A/Experimental source: serovar 3; biovar 1
 A/Genetics:
 A:Gene: pepf-2; U0521
 A:Genetic code: SGC3

Query Match 3.5%; Score 136; DB 2; Length 611;
 Best Local Similarity 19.5%; Pred. No. 0.13;
 Matches 119; Conservative 100; Mismatches 256; Indels 134; Gaps 27;

QY 31 NYNTITTEENVONMNNAGDKWSAFK-ESSTLAQMYPLQEIQNLTVK----- 76
 DB 69 NYVSNKLTQNL-IDNEMLAMSOKIEHQHRAKIF--INFENIAIKNNDLINSLKNS 124
 QY 77 -----LQDALQOQSSVLSDESKRLNTILNTMT-----YSTG 112
 DB 125 LKRYOLEYELMWYERKHIIEQOQKVYTAISRESSFGDIFVLLDSMDQYDGINYKQ 184
 QY 113 KVCNPNPOELLERGLNEIMANSLDYNERLWAMSRSSEVKOLRPLYEYVVKNNM 172
 DB 185 KVCFKK--QTDLYVAKSMDRALRKSAVSHRAIYDLNRTSKL---LYEY-VKQNEL 238
 QY 173 ARAHHEDYGDWKGDEYVNGVDYDSGOLIEDVHTFEIKPLYLEHAAVVR--AKL 230
 DB 239 AKLHNFKDY-----ISADAFSDKVDKNF-----INHYYQTKFAFG 275
 QY 231 MAAVPIYISPIGLPAHLGLDMGKRWTLISLTPFGKPNIDVTDA--VQAMDQRI 289
 DB 276 INRYTYRT-----LELKOQYOLTKVEPWDMKMLDIDKKNMESISAKML 320
 QY 290 FKAKEFPVSVGLPNTQGFENSMULTDGNVOKAVCHPTAMDLDGDRILMCTKVTMD 349
 DB 321 TLEALLLGSSEYINVOAFNEQIMWMMNNKISAVSISNTKGLDKFIILMNDETYN 380
 QY 350 DELLAAHMGHIQYDAAVAQPELFRNGANEGFHEAVGEIMLSAATPHLKSIGLSPD 409
 DB 381 SILTLVHELGHSHV--TYFA-----NOSQEVYNEEYFAELASTINELIMNYHLK-K 431
 QY 410 FQDNDETEINFLKQALTYGILPFTYMLEKRWMTYFKEIPKQDMKKWKKREYVG 468
 DB 432 YENDDLMLRLYLIDEMISGFATTTTCAIFSNFEMVA-----NEWINQGEFSWKKIYL 484
 QY 469 VVEPVPHDET-----YCDPASLPHVSDYEFIRYRTTLTFOFOF--EALQQA 516
 DB 485 AVEIEHNDITGYKNNKISKIDEANALLIINPH---FYGNFVYKVVIGQIC----- 536
 QY 517 HEGPLKCDI--SNSTAGOKLFNMLRLGRS--EFTLALENVYGAKNMVRPLNTEPLP 574
 DB 537 --GLINAIIRFNKNNAKAKERYCFEKFSGSLSP--LETINIILDIK--INENDWAEVNIIF 591
 QY 575 -TWLQDNK 582
 DB 592 NSMIDYIK 600

RESULT 20

D69943

carboxypeptidase homolog ypwA - *Bacillus subtilis*C/Species: *Bacillus subtilis*

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C/Accession: D69943

R/Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
 Nature 390, 249-256, 1997
 A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal
 lech, J.; Harwood, C.R.; Henuat, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
 Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
 A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
 y, M.; Ogawa, K.; Ogilwara, A.; Oudaga, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portele
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scani
 A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
 kouchi, M.; Takamashi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiya
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
 A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*
 A/Reference number: A69580; MUID:98044033; PMID:9384377
 A/Accession: D69943
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-501 <KUN>
 A/Cross-references: GB:299115; GB:AL009126; NID:g2634478; PIDN:CAB14125.1; PID:g26346
 A/Experimental source: strain 168
 C/Genetics:
 A:Gene: ypwA
 C:Superfamily: Thermus aquaticus carboxypeptidase Taq

Query Match 3.5%; Score 135; DB 2; Length 501;
 Best Local Similarity 20.9%; Pred. No. 0.11;
 Matches 98; Conservative 65; Mismatches 164; Indels 142; Gaps 26;

QY 162 YEEYVVL--KNDMA--RANHEDYG-----DVRQDYEYNGVDYDS 200
 DB 103 YKEVYLLSKETAMEAKGSDSLSPYLEOLIEFKRITW-----GQ-- 150
 QY 201 RGQLEDVHTFEIKPLYLEHAAVVRRAKLNAAVYSISPICLPAHLGLDMGKRWNTL 260
 DB 151 -----EHYDALDLDFE--PGVYKVLVD-----QLPAEL 177
 QY 261 YSLVTF-----GQRNIDVTDAVDQADQIRFEAKKFSVSGLPMTQGFENS 313
 DB 178 KEALIPVKQYVYASGNKP-----DTSITRAFPKCKELSLYFLDELGYDF-- 224
 QY 314 MULDGNVOKAVCHPTAMDLDGKDRILMCTKVTMDPLTA-----HHEMGHIQYD--MAY 367
 DB 225 ---DGRIDEYV--HFAITLNRGDYV--TTRYDEKDFRIALFTGTHCGAITEQINDE 278
 QY 368 AAQPELFRNGANEGFHEA-----VGE-----TMSAATPHLKSIGLSPDF- 410
 DB 279 ALSGINLSDGASMGHESQSLFEYENFGRNKHFWPTPKKIQEASVQFNDISL--DDFV 336
 QY 411 QEDNTEINFLKQA--ITVIGILPFTYMLEKRWMTYFKEIPKQDMKKWKKREYVG 468
 DB 337 RAINESKPSFIRVEADELTPHLIIRYEIK--AIFSNVSEVDLPILNOKRYODYL 393
 QY 469 VVEPVPHDETCDPASLPHVSDYS--FIRYRTTLTFOFOFALCO-----AAKHG 521
 DB 394 I---TPQDASITLQDVMAGAGDGYFPSYALGTYTAOLQKIMEDILPEEDALLERGF 450
 QY 532 HKCDISNSTAGOKLFNMLRL--GKSEPTLALENVYGAKNMVRPLNTEPLN 569
 DB 451 HPIK-----QWLEKVIHIGKRRKPLDIKDQATG--EELNVRILDY 490

Search completed: May 26, 2003, 17:57:42

Job time: 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 16:53:03 ; Search time 21 Seconds

(without alignments)
1422.046 million cell updates/sec

Title: US-09-978-385-2_COPY_19_738

Perfect score: 3869
Sequence: 1 STIEEQAKTFLDKFNHEAD.....DNSLEFLGIQPTLGPNNPP 720

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1337	34.6	732	1	ACET_HUMAN
2	1337	34.6	1306	1	ACET_HUMAN
3	1334	34.5	732	1	ACET_MOUSE
4	1334	34.5	1312	1	ACE_MOUSE
5	1312	33.9	1193	1	ACE_CHICK
6	1310	33.9	1313	1	ACE_RAT
7	1283	33.2	737	1	ACET_RABIT
8	1283	33.2	1310	1	ACET_RABIT
9	1073	27.3	611	1	ACE_DROME
10	1054.5	27.3	501	1	ACE_HARE
11	135	3.5	501	1	YPMW_BACSU
12	135	3.2	627	1	GLGB_BACSU
13	121	3.1	1034	1	BGAL_BACME
14	119	3.1	986	1	EPIT_STAEF
15	115.5	3.0	3433	1	UTRO_HUMAN
16	115	3.0	950	1	Y511_RICPR
17	113.5	2.9	3655	1	YAMB_SCHPO
18	113	2.9	6669	1	NEBU_HUMAN
19	112	2.9	663	1	WTR_CABEL
20	112	2.9	3911	1	AKA9_HUMAN
21	111	2.9	4540	1	DYHC_PARTE
22	109	2.8	906	1	Y010_CLOAB
23	108.5	2.8	773	1	DNA2_SCHPO
24	107.5	2.8	1398	1	DNA2_SCHPO
25	107.5	2.8	1689	1	RPA1_SCHPO
26	107	2.8	736	1	VPA_SCHPO
27	107	2.8	1284	1	CMFB_SCHPO
28	106.5	2.8	1084	1	XPOL_YEAST
29	106.5	2.8	1312	1	RASD_YEAST
30	105.5	2.7	645	1	REP_BUGAT
31	105.5	2.7	1002	1	POI_HY104
32	105	2.7	1279	1	BCHH_CHLVI
33	105	2.7	2167	1	YCS2_YEAST

Not found

ALIGNMENTS

RESULT 1	ACET_HUMAN	STANDARD:	PRT:	732 AA.
ID	ACET_HUMAN			
DT	P22966;			
DT	01-AUG-1991 (Rel. 19, Created)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Angiotensin-converting enzyme, testis-specific isoform precursor (EC 3.4.15.1) (ACE-T) (Dipeptidyl carboxypeptidase I) (Kinase II).			
GN	DCPI OR DCP OR ACE.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90046671; PubMed=2554286;			
RA	Ehlers M.R.W., Fox E.A., Strydom D.J., Riordan J.F.;			
RT	"Molecular cloning of human testicular angiotensin-converting enzyme: the testis isozyme is identical to the C-terminal half of endothelial angiotensin-converting enzyme.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 86:7741-7745(1989).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=8938720; PubMed=2547653;			
RA	Latlion A.L., Soudrier F., Allegrini J., Hubert C., Corvol P.,			
RT	Alhenc-Gelas F.;			
RT	"The testicular transcript of the angiotensin I-converting enzyme encodes for the ancestral, non-duplicated form of the enzyme.";			
RT	FEBS Lett. 252:99-104(1989).			
RL	[3]			
RP	SEQUENCE FROM N.A., AND VARIANTS P-32; G-49 AND S-712.			
RX	MEDLINE=99251580; PubMed=10319862;			
RA	Rieder M.J., Taylor S.L., Clark A.G., Nickerson D.A.;			
RT	"Sequence variation in the human angiotensin converting enzyme.";			
RT	Nat. Genet. 22:59-62(1999).			
RL	[4]			
RN	ZINC-BINDING.			
RX	MEDLINE=91308093; PubMed=1649623;			
RA	Ehlers M.R., Riordan J.F.;			
RT	"Angiotensin-converting enzyme: zinc- and inhibitor-binding stoichiometries of the somatic and testis isozymes.";			
RT	Biochemistry 30:7118-7126(1991).			
RL	[5]			
RP	FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF THE C-TERMINAL DIPEPTIDE.			
RX	THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN. ALSO ABLE TO INACTIVATE BRADYKININ, A POTENT VASODILATOR.			
CC	CATALYTIC ACTIVITY: Release of a C-terminal dipeptide, oligopeptide-1-Xaa-Xbb, when Xaa is not pro, and Xbb is neither Asp nor Glu. Converts angiotensin I to angiotensin II.			
CC	- CORFACTOR: BINDS 1 ZINC ION.			
CC	- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	- ALTERNATIVE PRODUCTS: 2 isoforms; testis-specific (shown here) and somatic (AC P12821); are produced by alternative splicing.			
CC	- TISSUE SPECIFICITY: SPERMATOCTES, ADULT TESTIS.			
CC	- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL			

[illegible]

RX MEDLINE-90318396; PubMed-2164636;
 RA Howard T.E., Shah S.-Y., Langford K.G., Martin B.M., Bernstein K.E.;
 RT "transcription of testicular angiotensin-converting enzyme (ACE) is
 RT initiated within the 12th intron of the somatic ACE gene.";
 RL Mol. Cell. Biol. 10:4294-4302(1990).
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN II TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-1-Xaa-Xbb, when Xaa is not pro, and Xbb is neither
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.
 CC -1- COFACTOR: BINDS 1 ZINC ION.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; testis-specific (shown here) and
 CC somatic (AC P09470); are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: SPERMATOCYTES, ADULT TESTIS.
 CC -1- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL
 CC -1- REGULATION BY ANDROGENS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M55333; AAA37149.1; -;
 DR EMBL: M61094; AAA37150.1; -;
 DR PIR: A35655; A35655.
 DR MEROPS: M02.004; -;
 DR MGD: MGI:87874; Ace.
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR000130; Zn_Metpeptidase.
 DR Pfam: PF01401; Peptidase_M2; 1.
 DR PRINTS: PR00791; Peptidase_M2.
 DR PRODOM: PD004184; Peptidase_M2; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR Hydroxylase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 KW Glycoprotein; Transmembrane; Testis; Signal; Alternative splicing.
 FT SIGNAL 1 31
 FT CHAIN 32 732
 FT PT ANGIOTENSIN-CONVERTING ENZYME,
 FT TESTIS-SPECIFIC ISOFORM,
 FT EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 32 684
 FT TRANSMEM 685 701
 FT DOMAIN 702 732
 FT METAL 413 413
 FT ACET SITE 414 414
 FT METAL 417 417
 FT METAL 441 441
 FT CARBOHYD 102 102
 FT CARBOHYD 120 120
 FT CARBOHYD 139 139
 FT CARBOHYD 185 185
 FT CARBOHYD 367 367
 FT CARBOHYD 616 616
 FT SEQUENCE 732 AA; 84047 MW; 16C817E7FBD09BD9 CRC64;
 SQ
 Query Match 34.5%; Score 1334; DB 1; Length 732;
 Best Local Similarity 42.6%; Pred. No. 2e-87;
 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;
 OY 2 TIEBAKTFLLKFNHEDLFYQSSLSMNTNTITEENQNMNADKSAFLKESTL 61
 DB 69 TDEAADRFVEYEDRTAQLVLEAEANWQNTITTEGSKILLEKSTESNHTLKGTG 128
 OY 62 AOMVPLQETONLVYLAQALQNGSSVLSSEKSKRLTTLTNTSTSTGKVCNPNPQ 121
 DB 139 AKTFVSNFQSSIKRIKIKONDRALVLPKLEEEYQIILLDETTYSLSNICYTG-- 186
 OY 122 ECLLEPGLNEMANSIDYNERLAWMSWSEVSKQPLPEYEVVLKEMARANYEDY 181

DB 187 TCMPLPDLTNMATSRYEELLMAMKSRDKVGRALLPPPKYVERSNKIATKGYTDA 246
 OY 182 GDYRGDEYVNCVDGYDYSRGOLIEVHEFEIEIKPLTEHIAVYRAKIMNAPYS-YISP 240
 DB 247 GDSMSLYESDLE-----QDEKLVLQELQPLTLMHAAVYRSLSHHYSEYINL 296
 OY 241 IGCPLAHLLGDWGFMTNLVSLVPGQKPNIDVTAMDQANDAOIRFEAKFEYVS 300
 DB 297 DQIPAHLLGNMAQTNISNIDVLPFPSPANIDATEAMIMQGMTPRIFEAQNFETSL 356
 OY 301 GLPNTOGFENSMITLDGNVOKACVPTADMLDGG--DEFITMKTVMDFLTAHHEMG 359
 DB 357 GLPVPPEFWKSMLEKPLDGEVCHSANDFYGADFRKQCTSVAMEDLVTAHHEMG 416
 OY 360 HIOYDMAAOPFLRNGANEHFAVEISLSNATPKHUKSIGLSLPDFQEDNEFEIN 419
 DB 417 HIQYFMQKNDLVETREGANPGFHEAIGDIMALSVSTKHLYSLNLSTE-GSGEYDIN 475
 OY 420 FLIKQALITVGLPTTYMLEKRMVVFGEIPIKQDMKMKEMKREIVGYVEPYPHDEY 479
 DB 476 FLMKALDKIARIPSYSLIDQWRWRFVFGSTIKENTYNDMMWSLRKLYGLCPVPRSGD 535
 OY 480 CPASLFHNSDYSTRYTRTYRILYQFOEALQAAKEGPHKCDISNSTEAGOKLFNM 539
 DB 536 FPDGSEFHPANVPVRYRVSPILQPFHEALCRAAGHTGPLHKDIIQSKAEAGKLADA 595
 OY 540 LRLGSEPTALLENVYAKNNVRPLNYEPFLTWLKDOK--NSFVGM-STDWSP 594
 DB 596 KRLGYSKFWPEAMKILITGQPNMSASAMMYEPLTEMLVTEHRRRGELTIGPEYVWAP 653
 RESULT 4
 ID ACE_MOUSE STANDARD; PRT; 1312 AA.
 AC P09470;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)
 DE (ACE) (dipeptidyl carboxypeptidase I) (kininase II).
 GN DCP1 OR ACE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP MEDLINE-89308599; PubMed-2545691;
 RA Bernstein K.E., Martin B.M., Edwards A.S., Bernstein E.A.;
 RT "Mouse angiotensin-converting enzyme is a protein composed of two
 RT homologous domains.";
 RL J. Biol. Chem. 264:11945-11951(1989).
 RN [2]
 RP SEQUENCE OF 1-332 FROM N.A. AND PARTIAL SEQUENCE.
 RA MEDLINE-88298730; PubMed-2841312;
 RA Bernstein K., Martin B.M., Bernstein E.A., Linton J., Striker L.,
 RA Striker G.;
 RT "The isolation of angiotensin-converting enzyme cDNA.";
 RL J. Biol. Chem. 263:11021-11024(1988).
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-1-Xaa-Xbb, when Xaa is not pro, and Xbb is neither
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.
 CC -1- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; somatic (shown here) and testis-
 CC specific (AC P22967); are produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: J04946; AAA37147.1; -
 DR EMBL: J04947; AAA37148.1; -
 DR EMBL: J03940; AAA37146.1; -
 DR PIR: A29220; A29220.
 DR PIR: A34171; A34171.
 DR MEROPS: M02.001; -
 DR MEROPS: M02.004; -
 DR MGD: MGI:87874; Ace.
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01401; Peptidase_M2; 2.
 DR PRINTS: PR00791; PEPIDPASEA.
 DR ProDom: PD004184; Peptidase_M2; 2.
 DR PROSITE: PS00142; ZINC_PROTEASE; 2.
 DR Hydroxylase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 KW Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.
 FT SIGNAL 1 34
 FT CHAIN 35 1312 ANGIOTENSIN-CONVERTING ENZYME, SOMATIC
 FT ISOPORM.
 FT DOMAIN 35 1264 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1265 1281 POTENTIAL.
 FT DOMAIN 1282 1312 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 232 588
 FT REPEAT 830 1186
 FT METAL 335 395 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 336 396 1 (BY SIMILARITY).
 FT METAL 339 393 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 933 993 ZINC 2 (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 994 994 2 (BY SIMILARITY).
 FT METAL 997 997 ZINC 2 (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 59 59 ZINC 2 (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 79 79 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 116 116 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 165 165 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 514 514 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 682 682 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 700 700 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 719 719 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 765 765 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 947 947 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1196 1196 N-LINKED (GLCNAC...) (POTENTIAL).
 FT VARIANT 568 568 T -> A.
 SQ SEQUENCE 1312 AA; 150947 MW; 9C13BB0529AD3755 CRC64;

Query Match 34.5%; Score 1334; DB 1; Length 1312;
 Best Local Similarity 42.6%; Pred. No. 4,5e-87;
 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

QY 2 TTEEDAKTFLDKFNHEADLEFYQSSLASMTNTNTNTEENVONMANNAGKWSAFLKEQSTL 61
 DB 649 TDEADAREVEYEDTAYVLLNEVFAEMQYNTNTTIGSKILLEKSTEVSHHTLKYGSTR 708
 QY 62 AOMYPLQEIQLTVLQLOAQNGSSVLSSEKSRKRLTILNTMTSTIYSGKVCNPDNPQ 121
 DB 709 AATFVSNFQNSIRIKTKLQNDRAVLPRELEEVQDILLDMETYSLSNICYTN-- 766
 QY 122 ECLLLEPGLEIMANSLDYNERLWAMESRSEVQKOLPLVEEYVVLAKENMARAHYDY 181
 DB 767 TCMPLPEPLTNMAMTSRKYBELLMAMKSMROKVGAILPEPKVYEFENKAKLNGYIDA 826
 QY 182 GDYMGDYEVNCGVDGYDYSGKOLLIDVHTTEELIKPLYLEHAAVYRAKLANAYS-YISP 240
 DB 827 GDSMSLYESDNLE-----ODLEKTKLOELPLYLNLHAYVRSRLHHYGESEYINL 876
 QY 241 ICSLPALHLLGDMWGRFWNTLVSLVVPFGQKPNIDVTDMANVQAMDAQIIFNEAEFFVS 300

DB 877 DGIPLAHLLGDMWGRFWNTLVSLVVPFGQKPNIDVTDMANVQAMDAQIIFNEAEFFVS 936
 QY 301 GLPMNTOGFENSMMLTPGNAVOKAHPRTAMDLGK-DFFILNCTKTYMDFTAHENG 359
 DB 937 GLTPVPEFWMKSKLELPTDREYVCHPSANDFNGNDPFRKCTSYNMEDLYAHENG 996
 QY 360 HIOYDAAVAAOPELLRNGANGFHEAVGEIMSLSAATPKHLKSTIGLSPDQDNTEIN 419
 DB 997 HIOYFMQYKIDLPYFRGANGFHEAIGDIALSVSTPKHLKSTIGLSPDQDNTEIN 1055
 QY 420 FLTKQALTYGTLPTFTMLEKRWVKGELPPDQNMKKMKREKRELVGVVEVPHDEY 479
 DB 1056 FLTKMALDKLAFPTFTSTLDQWRWVKGELPPDQNMKKMKREKRELVGVVEVPHDEY 1115
 QY 480 CDPASLFHVSNDYSFIHYRTLYFOFOBALCOAKHEPLHKCDISNSTEAGOKLFNM 539
 DB 1116 FDPGSKFHPANVYVYFVSFIQFQFHEALCAACHTPTPLKCDIYOSKEAGKLLADA 1175
 QY 540 LRLCKSEPTLALENVGAKNNVRLPLNPEPLTWLKDQNK--NSFVGW-STDWSP 594
 DB 1176 MKLGSKPWEPMKRLITGPNMSSAMNNYFKPLTEWLVTENRHGETLGMPEYNAP 1233

RESULT 5

ACE-CHICK STANDARD; PRT: 1193 AA.

ID ACE-CHICK STANDARD; PRT: 1193 AA.
 AC 010751;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Angiotensin-converting enzyme (EC 3.4.15.1) (Dipeptidyl
 DE carboxypeptidase I) (Kininase II) (Fragment).
 GN DCP1 OR ACE.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OC NCBI_TaxID=9031;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=lung;
 RA MEDLINE=95110342; PubMed=7811282;
 RA Esther C.R., Thomas K.E., Bernstein K.E.;
 RT "Chicken lacks the testis specific isozyme of angiotensin converting
 RT enzyme found in mammals";
 RL Biochem. Biophys. Res. Commun. 205:1916-1921(1994).
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-I-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.
 CC -1- COFACTOR: BINDS 2 ZINC IONS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

KW Hydroxylase: Metalloprotease: Carboxypeptidase; Zinc; Dipeptidase;
 KW Glycoprotein; Transmembrane; Repeat.
 FT NON_TER 1
 FT DOMAIN <1 1156 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1157 1173 POTENTIAL.
 FT DOMAIN 1174 1193 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 125 481
 FT REPEAT 723 1079
 FT METAL 288 288 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 289 289 1 (BY SIMILARITY).
 FT METAL 292 292 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 886 886 ZINC 2 (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 887 887 2 (BY SIMILARITY).
 FT METAL 890 890 ZINC 2 (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 485 485 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 555 555 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 575 575 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 658 658 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1089 1089 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1193 AA; 137820 MW; 954472A1B8A471C7 CRC64;

Query Match 33.9%; Score 1312; DB 1; Length 1193;
 Best Local Similarity 40.4%; Pred. No. 1,4e-85;

Matches 264; Conservative 110; Mismatches 243; Indels 36; Gaps 10;

OY 4 EEOAKTELDKFNHEADLFYQSSIASNNVTNTTEENVOANNAGDKWSAFLEKQSTLAQ 63
 DB 544 EAOAKTELDKFNHEADLFYQSSIASNNVTNTTEENVOANNAGDKWSAFLEKQSTLAQ 63
 OY 64 MYRQELONLVTKQLQALQONSSVLSSEKSKRLNLTMTGIVSGVCPND---P 120
 DB 604 QFDPSPDFQDVTYVTKLKLVSLEALPDELEKYNLTLSDMETYSVAVCCRBNTFHP 663
 OY 121 QECLEFGLNEIANSIDYNERLAWESNRSEVGKOLRPLYEYVLKKNEMARAHYED 180
 DB 664 -----LDPLDILATSRDYNELLFAKGMWDASGAIKIKRYVELSKAAVNLGYTD 718
 OY 181 YGDIWRGDYEVNGVDYDSRGQLEDEVENTFEIEIKLYEHLHAYVRAKIMAY-PSYIS 239
 DB 719 NGATWRSLEYETPE-----EDLERLYLOPLLYLNLHAYVRRALYNNKYGAHIS 768
 OY 240 PIGGLPAHLIGDMGGRFTNLYSLTFPGQKPNIDVDADVADQAMDQRIFKAEKFTVS 239
 DB 769 LKGPTRPAHLIGDMGGRFTNLYSLTFPGQKPNIDVDADVADQAMDQRIFKAEKFTVS 239
 OY 300 VGLPNTQGFENSLMTDPGVNOKAVCHPTAWDL-GKGDPRILMCTKYVMDDEFLAHHEM 358
 DB 829 IGLIPMPQEFWDKSMIEKPADGRVYCHASAMDPYNNKDKRIRICQCTVYNNDDILYVHEM 888
 OY 359 GHIOYMAVYAAQPLPLNNGANEHVEGHEISAAFPYHLKISGLISDFQDNETET 418
 DB 889 GHVYFLQYMDQPSIFSDGANPGHEHIGDYMALYSTPRHLHSINLD--QVTEENESRDI 947
 OY 419 NELLKQALTYVGLTFYVYMLEKRMVYFKGEIPRDOMKRWMEKREIVAVVEVPDEET 478
 DB 948 NYLMSIALDKTAFLPFGYLDQMWKVFDRIKEDENYQOQWMLNRLYQGLCPVPSSED 1007
 OY 479 YCDPASLFEVNSDYSTRYTRTLQYFOFOALCOAHKHEGFLKDCISSTFAGQKLFN 538
 DB 1008 DFDDGAKFHPANPYRIRYFVYIQFOHALCKAAGHGPLHTOYISKRAKGLGD 1067
 OY 539 MLRIKSEPTWLALENVGAKNANVRLYFEPFLTWL--KQONKSFVGM-STDNSTPY 595
 DB 1068 AMKIGFSPWPEAMQITGQPNMSAEALMSYFEPFLTWL--KQONKSFVGM-STDNSTPY 595
 OY 596 ADGSIKVRISLKSALG-----DKAYEMNDENYLFRRSSVAVAMRQYELKVR 641

DB 1128 AVTEHPAHTDADFLGMSVGTNRQATAGAM-----VLALALVELTISFLGVR 1175
 RESULT 6
 ACCE_RAT ACCE_RAT STANDARD; PRT; 1313 AA.
 AC P47820;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)
 DE (ACE) (dipeptidyl carboxypeptidase I) (kininase II).
 GN DCP1 OR ACE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=94121658; PubMed=8292044;
 RA Koike G., Krieger J.E., Jacob H.J., Mukoyama M., Pratt R.E.,
 RA Dzuu V.J.;
 RT "Angiotensin converting enzyme and genetic hypertension: cloning of
 RT rat cDNAs and characterization of the enzyme."
 RL Biochem. Biophys. Res. Commun. 198;380-386(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LEW/N; TISSUE=Lung;
 RA Jafarian-Tehrani M., Listwak S., Barrientos R.M., Michaud A.,
 RA Corvol P., Sternberg E.M.;
 RT "Characterization of a missense mutation in the angiotensin
 RT 1-converting enzyme cDNA in exudative inflammation resistant F344/N
 RT rats."
 RL Submitted (NOV-1999) to the EMBL/Genbank/CDN databases.
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU. THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-1-xaa-xbb, when xaa is not pro, and xbb is neither
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.
 CC -1- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; somatic (shown here) and testis-
 CC specific; are produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: U03708; AAA82110.1; -
 DR EMBL: U03734; AAA82111.1; -
 DR EMBL: AF201332; AAG35597.1; -
 DR MEROPS: M02.004; -
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01401; Peptidase_M2; 2.
 DR PRINTS: PR00791; PEPTIDASEA.
 DR ProDom: PD004184; PEPTIDASE_M2; 2.
 DR PROSITE: PS00142; ZINC_PROTEASE; 2.
 KW Hydroxylase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 KW Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.
 FT SIGNAL 1 35
 FT CHAIN 36 1313
 FT DOMAIN 36 1265
 FT TRANSMEM 1266 1282
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.

FT	DOMAIN	1283	1313	CYTOPLASMIC (POTENTIAL).
FT	REPEAT	233	589	
FT	METAL	396	1187	ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE	397	397	1 (BY SIMILARITY).
FT	METAL	400	400	ZINC 1 (CATALYTIC) (BY SIMILARITY).
FT	METAL	994	994	ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE	995	995	2 (BY SIMILARITY).
FT	METAL	998	998	ZINC 2 (CATALYTIC) (BY SIMILARITY).
FT	CARBOHYD	44	44	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	60	60	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	80	80	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	117	117	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	152	152	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	166	166	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	324	324	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	515	515	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	683	683	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	701	701	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	720	720	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	766	766	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	948	948	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1197	1197	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	207	207	R -> K.
SO	SEQUENCE	1313 AA;	150907 MM;	8CB5D0015F129591 CRC64;
Query Match		33.9%;	Score 1310;	DB 1;
Best Local Similarity		42.0%;	Pred. No.2.3e-85;	
Matches 251;	Conservative 115;	MissMatches 214;	Indels 18;	Gaps 7;
QY	2	TIERQAKFLDKFNHEADLEFYQSSLSAMNTNTTEENVOGNMNAAGKMSAFLEKQSTL	61	
DB	650	TDEKANKFEVEYDITAKVLNNEFAEAMWHNTTIGSKILLQKNKEVSHHTLKYGTW	709	
QY	62	AQMTPLQEIQLIYKALQALQOQSSVLSDEKSKRLTITNTMSTYSTGVCVCPDNPQ	121	
DB	710	AKTQDVSNFQNSTIKRIKKYQVNDRAVLPRNELEBYNQILDMMETTYTSVANVCTNG--	767	
QY	122	ECLLEPGLNETMANSLDYNERLWAMESRSEVQKQRLPYEEVYVLKNEMARANHYEDY	181	
DB	768	TCLSLPEPLTINIMTSRKYEEELLTWKMSRQKVRGAILPPFPKYDVFENKILANGYSDA	827	
QY	182	GDVNRGDEVNGVGYDYSRQGLLEDVHTTEERKELYEHLHAYVRATLMAYPS-YISP	240	
DB	828	GDSMSSSTESDLE-----QDEKLVQELQPLYLNLHAYVRSLSHRYGSEYTNL	877	
QY	241	IGCLPAILHLLDMGGRFTNTLXSLVYFQSKRINIDTDMVQDAQORLFEKEKFEVSV	300	
DB	878	DGPIPAHLILGMMMAQTMNTNIDVAPFSPASISIDTEAMIKQGTPKRIKFAKDAFTSL	937	
QY	301	GLPNKTOGFWENSMLTDGPNVQAKVCHPTANDLGK--DEFLIMCTKYTWDFELTAHHENG	359	
DB	938	GLLPVPEFPMKSMLEKPTDGRREVYCHASAMDFYNGKDFRIKQCTSVNMEELVIAHHENG	997	
QY	360	HIQYDMAAAPPELLRANANGFHNVAVEIMSLSATYKHLKSLGSLSPDQEDNETEIN	419	
DB	998	HIQYFMQKDLPVYFREGANGFHEAIDVLAIVSTYKHLHSLMLSS--GSGYEHDIN	1056	
QY	420	FLFKQALITVGLPTVTYMLEKRMVWFGEGLPKDDMKMMKMKREIYGVVPHDETY	479	
DB	1057	FLMKALDLKIAFIPFSYLLIDQRMKRVFPGSITIKENTNDEWMSLRKLYGICLCPVPRSGQD	1111	
QY	480	CDPASLHFVNSDYSEIRYRYTRLVQFOEALCOAKAEGBPLHKDISNSTEAGOKRTNM	539	
DB	1117	FDPGSKFHVAPNVPIRYFISPIIDQFHEALCRAANGHTGPLYCKDITYOSKRAGLTDA	1177	
QY	540	LRLKSESWTALALENVGAKMANNRPLANTPEPLTWLKDQK--NSYVGW--STWSP	594	
DB	1177	MLGYSKQMPAMKLTITGQPNMSASAINMYEKPLTEMLVTENRRRGELTGWPEYTWTP	1234	

ID	ACET	RABIT	STANDARD;	PRN;	737 AA.
DT	01-AUG-1991	(Rel. 19, Created)			
DT	01-AUG-1991	(Rel. 19, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Angiotensin-converting enzyme, testis-specific isoform precursor				
DE	(EC 3.4.15.1) (ACE-II) (Dipeptidyl carboxypeptidase I) (kininase II).				
GN	DCPI OR ACE.				
OS	Oryctolagus cuniculus (Rabbit).				
OC	Eukaryota; Metacoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.				
OX	NCBI_Taxid:9986;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-New Zealand white; TISSUE-Testis;				
RA	Kunlar R.S., Kusari J., Roy S.N., Soffer R.L., Sen G.C.;				
RA	Kunlar R.S., Kusari J., Roy S.N., Soffer R.L., Sen G.C.;				
RT	"Structure of testicular angiotensin-converting enzyme. A segmental				
RT	mosaic isozyme".				
RL	J. Biol. Chem. 264:16754-16758(1989).				
RN	[2]				
RP	SEQUENCE OF 1-117 FROM N.A.				
RC	MEDLINE=91139683; PubMed=1847288;				
RA	Kumar R.S., Thekkumkara T.J., Sen G.C.;				
RA	"The mRNAs encoding the two angiotensin-converting isozymes are				
RT	transcribed from the same gene by a tissue-specific choice of				
RT	alternative transcription initiation sites.";				
RL	J. Biol. Chem. 266:3854-3862(1991).				
CC	-1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF				
CC	THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE				
CC	VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.				
CC	-1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,				
CC	oligopeptide- -xaa-xbb, when xaa is not pro, and xbb is neither				
CC	asp nor glu. Converts angiotensin I to angiotensin II.				
CC	-1- COFACTOR: BINDS 1 ZINC ION.				
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	-1- ALTERNATIVE PRODUCTS: 2 isoforms, testis-specific (shown here) and				
CC	somatic (AC P12822); are produced by alternative splicing.				
CC	-1- TISSUE SPECIFICITY: SPERMATOCTYES, ADULT TESTIS.				
CC	-1- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL				
CC	REGULATION BY ANDROGENS.				
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sdb.ch/announce/				
CC	or send an email to license@isb-sdb.ch).				
CC	-----				
DR	EMBL; J05041; AAA31153.1;				
DR	EMBL; M58580; AAA31152.1;				
DR	PIR; A34402; A34402.				
DR	MEROPS; M02.004; --				
DR	InterPro; IPR001548; Peptidase_M2.				
DR	InterPro; IPR000130; Zn_MTpeptidase.				
DR	Pfam; PF01401; Peptidase_M2; 1.				
DR	PRINTS; PR00791; PEPTIDPASEA.				
DR	ProDom; PD004184; Peptidase_M2; 1.				
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.				
RW	Glycolipase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;				
RW	Hydrolyase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;				
RW	Glycoprotein; Transmembrane; Testis; Signal; Alternative splicing.				
FT	SIGNAL	1	32		
FT	CHAIN	33	737		
FT					
FT	DOMAIN	33	690		
FT	TRANSMM	691	707		
FT	DOMAIN	708	737		
FT	METAL	419	419		
FT	ACT SITE	420	420		
FT	METAL	423	423		
FT	METAL	447	447		

DE carboxypeptidase I) (Kininase II).
GN ANCE OR PACE.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX Cornell M.J., Williams T.A., Lamango N.S., Coates D., Corvol P.,
RA Soudier F., Hohseil J., Lebrach H., Isaac R.E.;
RT "Cloning and expression of an evolutionary conserved single-domain
angiotensin converting enzyme from Drosophila melanogaster.",
RL J. Biol. Chem. 270:13613-13619(1995).
RN [2]
RP REVISIONS.
RA Cornell M.J.;
RP Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RC SEQUENCE FROM N.A.
RC STRAIN-Canton-S;
RX MEDLINE=96028519; PubMed=7547464;
RA Tatek K., Cai H., Ip Y.T., Levine M.;
RT "Race: a Drosophila homologue of the angiotensin converting enzyme.";
RL Mech. Dev. 51:157-168(1995).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE CONTRACTIONS OF THE HEART, GUT
AND TESTES.
CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
oligopeptide-I-xaa-xbb, when xaa is not pro, and xbb is neither
asp nor glu. Converts angiotensin I to angiotensin II.
CC -1- COFACTOR: BINDS 1 ZINC ION.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE AMNIOSESOSA DURING GERM BAND
ELONGATION, SHORTENING AND HEART MORPHOGENESIS. EXPRESSED IN
MIDGUT THROUGHOUT EMBRYOGENESIS
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its use by
CC the non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL, U25344; AACB02171.1; -.
CC EMBL, U34599; AAC46902.1; -.
DR MEROPS: M02.003; -.
DR FLYBase: FBgn0012037; Ance.
DR InterPro: IPR001548; Peptidase_M2.
DR InterPro: IPR000130; Zn_MTpeptidase.
DR Pfam: PF01401; Peptidase_M2; 1.
DR PRINTS: PR00791; PEPDIPFASA.
DR ProDom: PD004184; Peptidase_M2; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolyase; Metallopeptidase; Carboxypeptidase; Zinc; Dipeptidase;
KW Glycoprotein; Signal.
FT SIGNAL 1
FT CHAIN 17
FT METAL 18 615
FT ACET SITE 367 368 368
FT METAL 371 371
FT CARBOHYD 53 53
FT CARBOHYD 196 196
FT CARBOHYD 311 311
FT CONFLICT 48 51
FT CONFLICT 141 141
FT CONFLICT 293 293
FT CONFLICT 486 486
FT CONFLICT 533 533
FT CONFLICT 547 547
FO SEQUENCE 615 AA; 71025 MW; 35830D5EAF33CFB CRC64;

Query Match 27.7%; Score 1073; DB 1; Length 615;
 Best Local Similarity 36.5%; Pred. No. 6, 2e-69;
 Matches 218; Conservative 120; Mismatches 238; Indels 22; Gaps 9;

4 EEOAKFLDKFNEHAEADLFYOSSLASMNNTTEENVOONNAGDKMSAFLEKQSTLQA 63
 22 EIOAKFLENLKELKRTNVEEAMAGSTNDENEKKEISAELEKFKKVAEDST 81
 64 MYLQETIONLVKLOLQAOQSSVLSDEKSKRLTITNTSTIYSTKVCNPNPQC 123
 82 KFWMSYSEDELRKQFKALTKGLYALPBDVAELDTLSMESNPAKVCYDYKSTGC 141
 124 -LLEPGLEIMANSIDYNERLWAMESRSEVGKOLRPLYEEVYVVKEMARANEHEDG 182
 142 DIALDPEIEVYSKSRDHEALYVWPEFDKAGTAVRSQPERVEYELNTAAKLNNTSGA 201
 183 DYWRGDEVNGVDYDSRGQLIEDVEHTEEIKPLYLEHLAVRAKLMNAV-PSYISPI 241
 202 EAWLDEYE-----DDTEFOQLEDI---FADIRPLYQOIHGIVRFLRKHYGDAAVSET 251
 242 GCLPAHLGDMGRFPTNLSLVTPRGOKPNIDVTAMDOQMDAQRIKKEAEKFPVSG 301
 252 GFLPMHLGMMWAOQMSIADIVSPPEKPLVDVSAEMEKQATPLKMGOGDFTSKN 311
 302 LPMNTOGFEMENSLTPGNAVOKAVCHPTAMDLGK-GDFRILMCTKYTMDFTLAHEMCH 360
 312 LTKLPDPMOKSTIEKPTDGRDLVCHASAMDFLIDVRLKQCTRTYQQLFVHHELG 371
 361 IYDAMVAAQPTLLRNGANGFHEAVGEIWSLSAATPKHLKSGLSPPFOEDNETIWF 420
 372 IYQFLYQHPFYRGANGFHEAVGDVLSVSPKHLKELGLK-DYVRDEARINQ 430
 421 LKQALTYVLTPTVWLEKRWMAVFGELIPKDOMKKWEMKREIVGVEPEPHETYC 480
 431 LEFLALDKIYFLPFAITMDKTYRMSLFRGEYDKANMCAFKLDESGIEPPVRSKDF 490
 481 DPASLFVNSDYFIRYRTTYLVOFOEALC-OAAKH-----EGPLHKDCISNSTEAG 534
 491 DAPAKYHISADVELRVLFIIOFQYKACIKAGQYDPAVDNELPDMNDIYGASARA 550
 535 KLRMLRLGSEPTLALENVGAKNNRPLNTEPEPLTWLAKDN--KNSVVGST 590
 551 AFHNMLSMGASKPMPALAEFNGERLMSKALAEYEPLEKVLAEENIKNNHIGWT 608

RESULT 10

ACE_HAEIE STANDARD; PRT: 611 AA.
 AC Q10715;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE angiotensin-converting enzyme precursor (EC 3.4.15.1) (Dipeptidyl
 carboxypeptidase I) (kininase II).
 GN ACE.
 OS Haemaphysa irritans exigua (Buffalo fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscoidae; Muscidae; Muscidae; Haematobia.
 NCBI_TaxID=34678;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96215437; PubMed=8647080;
 RA Wajfelds G.L., Fitzgerald C., Gough J., Riding G.A., Elvin C.,
 Kemp D.J., Willadsen P.;
 RT "Cloning and characterization of angiotensin-converting enzyme from
 the dipteran species, Haematobia irritans exigua, and its expression
 in the maturing male reproductive system".
 RL Eur. J. Biochem. 237:414-423(1996).
 CC -1- FUNCTION: INVOLVED IN THE SPECIFIC MATURATION OR DEGRADATION OF A
 NUMBER OF BIOACTIVE PEPTIDES.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,

oligopeptide-1-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
 Asp nor Glu. Converts angiotensin I to angiotensin II.
 -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
 -1- TISSUE SPECIFICITY: EXPRESSED IN THE COMPOUND GANGLION AND IN THE
 POSTERIOR REGION OF THE MIDGUT.
 -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

EMBL: L43965; AAA70427.1; -.
 DR MEROPS; M02.003; -.
 DR InterPro; IPR001548; Peptidase_M2.
 DR InterPro; IPR000130; Zn_MTPeptide.
 DR Pfam; PF01401; Peptidase_M2; 1.
 DR PRINTS; PR00791; PEPDIPASA.
 DR Prodom; PD004184; Peptidase_M2; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR Hydrolase; Metalloprotease; Zinc; Dipeptidase;
 KW Glycoprotein; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 611
 FT METAL 367 367
 FT ACT_SITE 368 368
 FT METAL 371 371
 FT CARBOHD 53 53
 FT CAROHD 196 196
 FT CAROHD 531 531
 SQ SEQUENCE 611 AA; 70505 MW; A43D6DF5A83EBCB53 CNG64;

Query Match 27.3%; Score 1054.5; DB 1; Length 611;
 Best Local Similarity 37.7%; Pred. No. 1, 3e-67;
 Matches 222; Conservative 105; Mismatches 239; Indels 23; Gaps 8;

1 STEEQAKT-FLDKFNEHAEADLFYOSSLASMNNTTEENVOONNAGDKMSAFLEKQ 59
 18 ARKEIVATEYVIONNKELAKHTNVEEAMAGSTNDENEKKEISAELEKFKKVAE 77
 60 TLAQVPIQETIONLVKLOLQAOQSSVLSDEKSKRLTITNTSTIYSTKVCNPNP 119
 78 KDIOKFNKRTYGSAPVRRQFSLKSTGYSALEADYALTELVLASMESNPAKVCYDY 137
 120 PQC-LLEPGLEIMANSIDYNERLWAMESRSEVGKOLRPLYEEVYVVKEMARANEH 178
 138 SAKCDLSIDPEIEEITTSRDEELKYTWQFYDKAGTPTSNEKVEYELNTRAKLNNE 197
 179 EDYGDYWRGDEVNGVDYDSRGQLIED-VEHTEEIKPLYLEHLAVRAKLMNAV-PS 236
 198 TGAEWLDEYE-----DATFEOQLAEIPEIDIKPLDYQHGIVRFLRKHYGDAAV 246
 237 YISPTGCLPAHLGDMGRFPTNLSLVTPRGOKPNIDVTAMDOQMDAQRIKKEAEK 296
 247 VVSKGFLPMHLGMMWAOQMSIADIVSPPEKPLVDVSAEMEKQATPLKMGOGDFT 306
 297 FVSVGLPNNTOGFEMENSLTPGNAVOKAVCHPTAMDLGK-GDFRILMCTKYTMDFT 355
 307 FQSMGKRLKPLQFEMKSTIEKPTDGRDLVCHASAMDFLIDVRLKQCTRTYQQLFV 366
 356 HEMGIQYDMVAAQPTLLRNGANGFHEAVGEIWSLSAATPKHLKSGLSPPFOEDNE 415
 367 HEMGHQYFLYQHPFYRGANGFHEAVGDVLSVSPKHLKELGLK-DYVRDEARINQ 425
 416 TEINFLKQALTYVLTPTVWLEKRWMAVFGELIPKDOMKKWEMKREIVGVEPEPH 475
 426 ARINQFLALDKIYFLPFAITMDKTYRMSLFRGEYDKANMCAFKLDESGIEPPV 485
 476 DETYCDPASLFVNSDYFIRYRTTYLVOFOEALCOAA-----KHEGDLHKDCIS 529

DB 486 TEKDFDAPAKYHSADVEYLRLVLSFIIOFOFKSACINCEYVNPOTETPLDNCIYGS 545
 OY 530 TEAGOKLFENMLRLCKSEPTLALENVVAKNNVRLPLNPEPLTWL 578
 DB 546 KEAGLFFENMLSLGASKRFPWPDALFANGERTMGKALAEYFEPRLWLE 594

RESULT 11
 ID YPWA_BACSU STANDARD: PRT; 501 AA.
 AC P50848;
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical metalloprotease yPWA (EC 3.4.24.-).
 GN YPWA.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / Marburg;
 RX MEDLINE=96349105; PubMed=8760912;
 RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
 RA Serror P.;
 RT "Sequence analysis of the Bacillus subtilis chromosome region between
 RT the sera and kag loci cloned in a yeast artificial chromosome";
 RL Microbiology 142:2005-2016(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Boloetin A., Borchert S.,
 RA Boursier L., Brans A., Brun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capiano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerion I.F., Cummins N.J., Daniel R.A.,
 RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson F.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasaahara Y., Klier-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kunita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle B., Potworlik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche B., Roche M., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpeira P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weltenegeger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis";
 RT Nature 390:249-256(1997).
 RL Nature 390:249-256(1997).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M32.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC

DR EMBL: L47838; AAB38482.1; -
 DR EMBL: L77246; AAB96610.1; -
 DR EMBL: Z99115; CAB14125.1; -
 DR MEROPS: M32.0PW; -
 DR Subtilist: BG11458; YPWA.
 DR InterPro: IPR001333; Peptidase_M32.
 DR InterPro: IPR001330; Zn_MTPeptide.
 DR Pfam: PF02074; Peptidase_M32; 1.
 DR PRINTS: PR00998; CROXYPTASE.
 DR PROSITE: PS00142; ZINC_PROTEASE.
 DR Hypothetical protein; Hydrolase; Metalloprotease; Zinc;
 KW Complete proteome.
 FT METAL 265 265 ZINC (CATALYTIC) (POTENTIAL).
 FT ACT_SITE 266 266 ZINC (CATALYTIC) (POTENTIAL).
 FT METAL 269 269 ZINC (CATALYTIC) (POTENTIAL).
 FT METAL 269 269 ZINC (CATALYTIC) (POTENTIAL).
 SQ SEQUENCE 501 AA; 58174 MW; A/489BBAF38F82 CAC64;

Query Match 3.5%; Score 135; DB 1; Length 501;
 Best Local Similarity 20.9%; Pred. No. 0.036;
 Matches 98; Conservative 65; Mismatches 164; Indels 142; Gaps 26;

OY 162 YEEYVVL--KNEMA--RANHEDYD-----DYRGDYEYNGVDGYDS 200
 DB 103 YKEYVILGSKAETAMEAKGSDFSLESPYLEQLEFNKREITW-----GIG-- 150
 OY 201 RGLIDVHTFEETKPLVEHLHAYVRKIMNAPYSISPGCLPAHLIGDMGRFTNL 260
 DB 151 -----BHPYDALDLE---PGYVAVLD-----QLFAEL 177
 OY 261 YSLVTF-----GKPNIDVTDAVDQMDAQRIEFAKEFVSVGLPNMTGFWENS 313
 DB 178 KEAIIPLVQVNTASGKRP-----DTSFTKAFPEKRELSLYELGEGYDF-- 224
 OY 314 MLTPGVQKAVCHPAPMGLGKDPRIIMCTVTMDPFLTA-----HHMGHLOYD--MAY 367
 DB 225 ---DGGRLDFTV--HPRATLNRGDVY--TRIDEKDFATLFGIHCGHAIYQNDI 278
 OY 368 AAOEFLRNGANGFHEA-----VGE-----IMSAATPKHLKSTIGLSDF- 410
 DB 279 ALSTNTSLSDASNGIHESGLFENLGNKHPWPKYKIQEASPVQFKDSTL--DQFV 336
 OY 411 QEDNETINFLKQA--LTVGLPFTYMLEKRMVFEGLPKQOMKMKWEMKREIVG 468
 DB 337 RAINESKSPRIYEADELTPYLIIRYELER--AIFSNESVEDLPSLMOKIQDYLG 393
 OY 469 VVEVPVPHDEYICOPASLFVSNYDS--FIRYRTLYLOFOEALCO-----AAKHEGL 521
 DB 394 I---TPQTDAGLIDYVHAGGDPGTPPSYALGMYTAAQLKQMLDELPFALLERGF 450
 OY 522 HKCDISSTENGOKLFENMLRL--GKSEPTLALENVVAKNNVRLPLNT 569
 DB 451 HPIK-----QWLTEKVIHGRKKRPDIIRDATG--EELNVRLIDY 490

RESULT 12
 ID GLGB_BACSU STANDARD: PRT; 627 AA.
 AC P39118;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.16) (Glycogen branching
 DE enzyme).
 GN GLGB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=94195107; PubMed=8145641;
 RA Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
 RT "Glycogen in Bacillus subtilis: molecular characterization of an

DR	InterPro: IPR004193; Isoamylase_N.
DR	Pfam: PF001028; alpha-amylase_1.
DR	Pfam: PF02922; Isoamylase_N; 1.
KW	Glycogen biosynthesis; Transferase; Glycosyltransferase;
KM	Complete proteome.
FT	ACT_SITE 309 309 BY SIMILARITY.
FT	ACT_SITE 352 352 BY SIMILARITY.
FT	ACT_SITE 420 420 BY SIMILARITY.
DE	SEQUENCE 627 AA; 73665 MW; 64B0A553B6767BA CRC64;
Query Match	3.28; Score 125; DB 1; Length 627;
Best Local Similarity	18.68; Pred. 0.25;
Matches 131; Conservative	86; Mismatches 225; Indels 264; Gaps 37
QY	31 NYNTNITIEVQVQNNNAGDKWSAFL-----KEGSTLAQMYPILOEIONLTVKLOALQON 85
DB	56 DFNSSGSEEHVMRYVNDNGIWTFLFIGIGEKER-----YKYEIYTN-----N 97
QY	86 GSSVLSADSKKRNLNTLNTWSTYSGKVQNPDPQECILLRPLNEIMAN--SLDYNER 143
DB	98 GEIRLRADP-----YAIYS-----EYRPNLTASLTYLE 125
QY	144 LMAMESRSFVGKOLRPLYEEVYVLKNEKRAHNEYDGYWRCGYEVNGVGYDYSRGQ 203
DB	126 GYSMDQOKNQKKOKATLYEKRPFI-----YELHGSWK-----KHSQGRHYSYKE 171
QY	204 LIEDV-----EHTFE--EIKPLVEHLHAYRAKLMAVPSYISPGCLPAHLDDMGGRF 256
DB	172 LSGTLPLRYIKKGFTIELLPYVE--HPYDRS-----WQYQ 205
QY	257 WTVLYSLIYVFGOKPIPIDVTADMYQA-----WDAQRIEKAERKFVSYGLPNMT 306
DB	206 GGGYGYEPTIRFG--PPHDLK-KVVECHQOQNGIILLDPGPHFKDAHGLYMGDEGP--- 259
QY	307 QGFWENSMITDPGQNVQAKVCHPTAVDLKSGDFRILMCTKYTMDFLTAH-----HEMGHIQ 362
DB	260 --LYEKEERDEN-----WLMGTANFDL--CKRPVHSFLISNALYMAEFYHI 303
QY	363 YDMAVAAQPEFLRNGANGEGH--EAVGELIMSLSAATPKHLKSLIGLSPDFQEDNETEINFL 421
DB	304 -----DGFYDVAANLILYPPNODERH-----TNPYAVDFL 333
QY	422 LKQALITVGLPFTYML--EKRMWVFGKELPKD-----QMKKWMEMKREIVGYVEPV 473
DB	334 KKLNTQMEAREPHVMMLIEDSTEMPQYTGAAVEEGSLGFHTKYNMGW--ANDVLKWIETP 390
QY	474 PHDEIYCPASLFIHVSNDYSFIRYRTYTVQFOQELCOAKHNEGFLHKDISNSTENG 533
DB	391 PEERHHC-----HQLISFSL-----YASEHFVLDPFSDDEVY-----G 425
QY	534 OKLNMRLFKGSEPTALTALENVGAKNNVRLPNT-----PEFLFTWIK 578
DB	426 KK--SLINKMGDYW-----QKPAQYRLLLGVTYVHPGKKLIFMGSEFQGFDEW-K 473
QY	579 DONKNSFGNSTDSPYADSI-----KVRISLSALGDAYEMND--NEYML 624
DB	474 DTEQ---LDNFFLDSFPNQKASVFTODLLRFYOKSKILYEHNDHRAQSFEMIDVHNDEQSI 530
QY	625 FRSSVAYAMROYELKVKNMQMLFGEEDVRYANCKPRISNFEPFAR 670
DB	531 F-SPIRYGQKH-----GEALVITICNETPPVYHQIDYGV 563

CC Bacteria: Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxId=1404;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA STRAIN-DSM 319;
 RA Streptococcus J.
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 galactose residues in beta-D-galactosides.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: AJ000733; CAA04267.1;
 DR HSP: P00722; 18GL.
 DR InterPro: IPR004200; Bgal_small1.C.
 DR InterPro: IPR004199; Bgal_small1.N.
 DR InterPro: IPR001649; GH_2.
 DR Pfam: PF00703; Glyco_hydro_2; 1.
 DR Pfam: PF02836; Glyco_hydro_2.C; 1.
 DR Pfam: PF02837; Glyco_hydro_2.N; 1.
 DR Pfam: PF02929; Bgal_small1.N; 1.
 DR Pfam: PF02930; Bgal_small1.C; 1.
 DR PRINTS: PR00132; GLHYDRLASE2.
 DR PROSITE: PS00719; GLYCOSYL_HYDROL_F2_1; 1.
 DR PROSITE: PS00608; GLYCOSYL_HYDROL_F2_2; 1.
 DR HydroLase: Glycosidase.
 FW ACT_SITE 481 481 PROTON DONOR (BY SIMILARITY).
 FT ACT_SITE 547 547 NUCLEOPHILE (BY SIMILARITY).
 FT ACT_SITE 1034 AA; 118673 MW; 38644C9A69415E9 CRC64;
 SQ SEQUENCE
 Query Match 3.1%; Score 121; DB 1; Length 1034;
 Best Local Similarity 19.2%; Pred. No. 0.99;
 Matches 104; Conservative 66; Mismatches 165; Indels 206; Gaps 30;
 OY 11 LDKRHEADLFYSSSL-----ASMYNNINTEENYNN--NAGDKSAFLKE 57
 DB 260 LNYEHDTQDLTFEYMLDANAQEVLPLOTNLSVSDQRTVSJLTHIKSAKSA----- 315
 OY 58 OSTLAQWYPLQEIQLN-TVKLQLALQNGSSVLSSEKSK--RLNTLLNTMSTIYSTGK 113
 DB 316 -----ESPMLTIVL---SLKNAAGSIITETSCKYGPRFEELKNGMLMT--NGK 359
 OY 114 -----VCPNDNPOECLELLEPGLNEIMANSIDYERLM----- 145
 DB 360 RIVLGNVRNHEPDSYKGRAGITREDMIDILMKOHNNINAVKTSHPYDMSVWYELCNEYG 419
 OY 146 -----AM-----ESMRSEYKOLRPLYEE----- 164
 DB 420 IYVIDETNLHGTITTYLOEGEQKAVPGSKPEKKNVLDRCRSKATERRKNPSTIIMSLG 479
 OY 165 -----YVVLK-NEMARANYEDYGDYMGDYEVNGVDGYDYSRGOLDEYHT 211
 DB 480 NESFGENGQMYTFEKKEDSTRLVHYE--GIFHHRDYDASDIESTMYKP--ADVE-- 532
 OY 212 FEELKPLVYELHAIYRAKAMAYPSYISIGCLPRLHGLGDMGRWTNLYSLTVPFGOKP 271
 DB 533 -----RYALMNPCKPYITL---CEYSHAMGNSCG---NLYKWEILFDQY 570
 OY 272 NID---VTDAMVDQADQRIKFEAE--KEFVSVG---LPMNTQGEWENSLTDPGNY 321
 DB 571 ILQGGITMD-----WKDALQATADGDSYLAAGDFDGTN-DGNFCNGILFADGTA 623
 OY 322 QKAV-----CH-PTAN--DLKGDRIILMCTKVTYMDLTLAHHEKGL-----QYDAY 367
 DB 624 SPRIAEVKKCYGQVVKWTAVDPAKGRFV-----QKHLFTNLINAYDFW 667

OY 368 AAQPLRLNGANEHVEHVEINSLSATPKHLSIGLLSPDQENETELFLKQALT 427
 DB 668 IYE-----KNG-----ELVERKASILNVPACDTDELTLSTPLVQENBID-EFVLTLSLR 716
 OY 428 I 428
 DB 717 L 717
 RESULT 14
 EPB: STAP
 ID EPB: STAP STANDARD: PRT: 986 AA.
 AC P30195;
 DT 01-APR-1993 (rel. 25, Created)
 DT 01-APR-1993 (rel. 25, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Epidermin biosynthesis protein epib.
 GN EPB.
 OS Staphylococcus epidermidis.
 OG Staphylococcus epidermidis.
 OC Bacteria: Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxId=1282;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-TU 3298 / DSM 3095;
 RX MEDLINE-92155237; PubMed-1740156;
 RA Schell N., Engelke G., Augustin J., Rosenstein R., Ungermann V.,
 RA Goetz F., Brilian K.-D.;
 RT Analysis of genes involved in the biosynthesis of lantibiotic
 RT epidermin.
 RT Eur. J. Biochem. 204:57-68(1992).
 CC -1- FUNCTION: INVOLVED IN THE POSTTRANSLATIONAL MODIFICATION OF THE
 CC LANTIBIOTIC EPIDERMIN.
 CC SUBCELLULAR LOCATION: POSSIBLY ASSOCIATED WITH, AND ANCHORED TO,
 CC THE CYTOPLASMIC SIDE OF THE MEMBRANE.
 CC -1- SIMILARITY: TO B. SUBTILIS SPAB AND L. LACTIS NISB.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: X62386; CAA44253.1;
 DR Plasmid.
 FW SEQUENCE 986 AA; 116708 MW; 343CE98526847850 CRC64;
 SQ SEQUENCE
 Query Match 3.1%; Score 119; DB 1; Length 986;
 Best Local Similarity 18.2%; Pred. No. 1.3; 247; Indels 310; Gaps 40;
 Matches 150; Conservative 115; Mismatches 115;
 OY 2 TIEQANTFLDKENH-----AEDL--FYOSSLASMYNNINTEENYNNNAGDKK 51
 DB 339 TIRNYHEFFMDKYFEDLVNKLKLLSDINGEYPKKXSYFSNNI----- 383
 OY 52 SAKLEOSTLAQWYPLQEIQLN-TVKLQLALQNGSSVLSSEKSKRLNTLLNTMSTIYST 111
 DB 384 -AFLEK-----YLAIONSHIEITENDYKNLEK-NNTVSKNA- 421
 OY 112 GKVQNPDPNPOECLELLEPGLNEIMANSIDYNERLAWESWSENGKOLRPLYEEYVLAKE 171
 DB 422 -----PVSTEL-----YSRIY-----GNSIKG-YEDFAVYSP 449
 OY 172 MARANHYEDYGDYMGDYEVNGVDGYDYSRGOLDEYHTFEELKPLVYELHAYVRAKLM 231
 DB 450 LGSFVAGATFRF--TGNNFNK-----KKNQLOREIVH-----HNNYMNENGL 491
 OY 232 NAYPSYISIGCLPRLHGLGDMGRWTNLYSLTVPFGOKPNIYTDAMVDQADQRIK 291
 DB 492 ELSQLEPGLNSRNVIILNN--NRINYCNLNLNP--KSDIDINDIFIGATFVKLYIS 546

```

OY 292 E--AEKFEVSGLPNMTG-----FWNSMTLDPGN-----VOKA 324
DB 547 EKHSRLVEVSNSEFNEFSGSELYKRLREISFEKTFIOPITBEGIDSLPECPRIYKNI
OY 325 VCHPFLADLGKGRITMCTKVTWDDFLTAHHEMGHIOYMAVA-----368
DB 607 ILKPAVTKINSEMFSE---TENMLNRAIT-REKWHIPDVIIAFSDNLLNLNDKHL 662
OY 369 -----AQPFLNCGANEGHFAVGEIMS-----LSAATPR-----HLK 401
DB 663 IILKELKHKGRIRILSEFINESNMERMEIIVPLKYKTSILKROSPFIIPKNKHNK 722
OY 402 ---SIGLSPDFQEDNTEINFLKQALITVGLPTVLEWRAVVEGELPKQWAKK 458
DB 723 DWMSIHLSIPKYOD-----NFIQDYLIPITELKKNVNFNFYKRED---EDFLK- 773
OY 459 WMEKREIVGVVEVPDEYTCPSALFHVSNDSPIRY-----TRLVYOFQFQALC 512
DB 774 -LRLRRE-----DEDY-----SQYSPFKMKDYCLLSNLSLYSTVDYVP 813
OY 513 OAAKHGEPGLHKDISSTEGAKLFLMMLLGKSEPTTLALENVGAKNNVRPLINTEP 572
DB 814 EYVRYGGPHYIEDIENFMYDSL--SINIQSE-FKIPKEPIVA---ISIDFLIDYLE- 866
OY 573 LFTWLDQNNKSEFVGNSTWSPYADQSIKVRISLKSALGDKAYEWNDNEMTLFSSVAYA 632
DB 867 -----INKSEKELLINNA--EDLYRND-----888
OY 633 MROYFLKVNQMILFEGEDVRYANAKPRISFEVYAPAKNSVDIIP-----678
DB 889 IREY-----KNLAKL-----TNPKNDYELLKKEPMLHEFLNKI 924
OY 679 -----RTEVEKAIKMSRSRINDAF---RLNDSLEFGIOP 711
DB 925 SILENLKTLQKSLVTSRSRIGSPIHRCN-----RIGGINP 962

RESULT 15
UTRO_HUMAN
ID UTRO_HUMAN STANDARD; PRT; 3433 AA.
AC P46939;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE utrophin (dystrophin-related protein 1) (DPR1) (DRP).
GN UTRN OR DMDL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RA MEDLINE=93096045; PubMed=1461283;
RA Tinsley J.M., Blake D.J., Roche A., Fairbrother U., Riss J.,
RA Byth B.C., Knight A.E., Kendrick-Jones J., Suthers G.K., Love D.R.,
RA Edwards Y.H., Davies K.E.;
RT "Primary structure of dystrophin-related protein.";
RL Nature 360:591-593(1992).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 144-261.
RA MEDLINE=99141377; PubMed=9887274;
RA Keep N.H., Norwood F.L.M., Moores C.A., Winder S.J.,
RA Kendrick-Jones J.;
RT "The 2.0-A structure of the second calponin homology domain from the
RT actin-binding region of the dystrophin homologue utrophin.";
RL J. Mol. Biol. 285:1257-1264(1999).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 31-256.
RA MEDLINE=20113481; PubMed=10647184;
RA Keep N.H., Winder S.J., Moores C.A., Walke S., Norwood F.L.M.,
RA Kendrick-Jones J.;
RT "Crystal structure of the actin-binding region of utrophin reveals a
RT head-to-tail dimer.";

```

```

RL Structure 7:1539-1546(1999).
CC -1- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE
CC PLASMA MEMBRANE (BY SIMILARITY TO DYSTROPHIN).
CC -1- SUBCELLULAR LOCATION: NEUROMUSCULAR JUNCTION.
CC -1- TISSUE SPECIFICITY: MUSCLE.
CC -1- SIMILARITY: STRONG, TO DYSTROPHIN.
CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,
CC ABP-120, ABP-180, OR BETA-FODRIN).
CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.
CC -1- SIMILARITY: CONTAINS 20 SPECTRIN REPEATS.
CC -1- SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: X69086; CAA8829.1; -.
DR PIR: S28381; S28381.
DR PDB: 1BHD; 16-FEB-99.
DR PDB: 1QAG; 01-JAN-00.
DR Genew; HGNC:12635; UTRN.
DR MIM; 128240; -.
DR InterPro: IPR001589; Actbind_actin.
DR InterPro: IPR001715; Calponin-like.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR003122; Tacth.
DR InterPro: IPR002349; WW.
DR InterPro: IPR001202; WW_Rsp5_WWP.
DR InterPro: IPR000433; Znf_Z2.
DR Pfam: PF00307; CH_2.
DR Pfam: PF00397; WW_1.
DR Pfam: PF00435; spectrin; 19.
DR Pfam: PF00569; Z2; 1.
DR PRINTS; PR00403; WMDOMAIN.
DR SMART; SM00033; CH_2.
DR SMART; SM00150; SPEC; 18.
DR SMART; SM00319; Tacth; 1.
DR SMART; SM00456; WW; 1.
DR SMART; SM00291; Znf_Z2; 1.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS00159; WW_DOMAIN_1; 1.
DR PROSITE; PS00020; WW_DOMAIN_2; 1.
DR PROSITE; PS01357; ZF_Z2_1; 1.
DR PROSITE; PS01357; ZF_Z2_2; 1.
KW Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;
KW Repeat; 3d-structure; Zinc-finger.
FT DOMAIN 1 246 ACTIN-BINDING.
FT DOMAIN 2 31 135 CH 1.
FT DOMAIN 3 150 252 CH 2.
FT REPEAT 253 308 SPECTRIN 1.
FT REPEAT 309 417 SPECTRIN 2.
FT REPEAT 418 526 SPECTRIN 3.
FT REPEAT 541 637 SPECTRIN 4.
FT REPEAT 687 798 SPECTRIN 5.
FT REPEAT 803 902 SPECTRIN 6.
FT REPEAT 1016 1083 SPECTRIN 7.
FT REPEAT 1125 1230 SPECTRIN 8.
FT REPEAT 1248 1334 SPECTRIN 9.
FT REPEAT 1432 1541 SPECTRIN 10.
FT REPEAT 1544 1649 SPECTRIN 11.
FT REPEAT 1652 1753 SPECTRIN 12.
FT REPEAT 1910 1968 SPECTRIN 13.
FT REPEAT 1976 2081 SPECTRIN 14.
FT REPEAT 2258 2333 SPECTRIN 15.
FT REPEAT 2399 2440 SPECTRIN 16.

```

FT REPEAT 2443 2556 SPECTRIN 17.
 FT REPEAT 2559 2636 SPECTRIN 18.
 FT REPEAT 2658 2688 SPECTRIN 19.
 FT REPEAT 2691 2797 SPECTRIN 20.
 FT DOMAIN 2812 2845 WM.
 FT ZN_FING 3064 3111 22-TYPE.
 SQ SEQUENCE 3433 AA: 394488 MW: 44885.5 EAE8DA09F8585B5 CRC64;
 Query Match 3.0%; Score 115.5; DB 1; Length 3433;
 Best Local Similarity 18.5%; Pred. No. 13; Indels 265; Gaps 36;
 Matches 139; Conservative 98; Mismatches 249;
 11 LKFNHEADL-----FYQSLASWNTNTTEENVOANNAGDKMSAFLEKQSTL 61
 2533 LDMNORNDIAKAKSASIRAHLEASAKENRRLMSLEIKLWLNKDEE---LKKOMPI 2588
 62 AQWPIQETONTLVKIQALQALQNGSSVSEDKSKRLNTILMTSTIYSGKVCNDPQ 121
 2589 GGDVPAIOLQYDHCALRELEKEKESVLAANDQARV-----FLADQPIEAPR 2639
 122 ECLLEPGL-----NEIMANSLDYNERLWAMES-----WRSEVQQLAPRYEE 164
 2640 RNLOSKETELPERBAQKIAKAMKOSSEVEK---WESINAVTSNMOKVDALKEKLDL 2696
 165 YVVLKNEARANTYEDYGYR--GDYEVNGVGYDYSNGQLEDEVEHFEETIKPLEYEH 222
 2697 QGADDDLDADKKEASVNRNGMKPVGDLLDLSLDH-----LEKIMAREEIAPI- 2745
 223 HAYYRAKLINAYFYISIPICLP---AHLGDMWGHFTNLTSLTVPGQKRNIDVYA 278
 2746 --NRKVTYNLDSQSLPDLHPSLKMSRQLDLDMR-W-KLQVSV-----DD 2790
 279 WVDQAMDQRIKFEAREKFEVSVGL-----PNTQGFV 310
 2791 RLKQLOEAFHDFEGSSQHFLSTVOLPMQNSISHNKPYIINQOTQTCDDHREKMELF- 2849
 311 ESMCLDPGNV-----OKAVC-----HPTANDLKGDF- 338
 2850 --QSLADLNAYFSAFRAIKIRLQKALCLDLELSTNEIFRKH---KLQNDQLLS 2903
 339 --RLMCTRVYMDFLTAHHEKHLQ-----YDMVAAQDFELLRNGANGCFH 383
 2904 VPDVINCILTYTDGLEOMHKLDVNPCLVDMCLMNLNVDTG-----RTKIRVQS 2955
 384 EAVGETLSLS-----AATPKHL---KSIGLSPD----- 409
 2956 LKLG-LMSLSKGLLEEKRYLFRKVAAGPTMCDQQLGLLDAIOLPROLGEVAAFGS 3014
 410 -----FOEDNETEINFLKQALTYGILPFTYMLEKMWAVFKGEIPRODMKKMW 460
 3015 NIEPSVRSCHQNNN-----KEPISVKERI-DMW 3042
 461 EMKREIVGVEPVPHDETCDPA---SLFHVSNDSFYRYRTYLYQFOBALQOAAKH 517
 3043 HLEQQA--VMLPVILHRAAEETAKHQAKCKICEPIVGFRRYSKLHFYD--VQOS- 3096
 518 EGPLKCDISNSTAGOKL-FNNMLRL-----GKS-EPWTLELVNVGAKNNMVR-PLLN 568
 3097 -----CFEFGSKTAAGKRLHYPVWEVCIPTTSGEDVADPLKYVILKNFRSKYFAKPRIG 3150
 569 YFEELFTMLKDONKNSFVNSTDMSPYADS 599
 3151 YL-PVQYVLEGDNIETPTLLISMWPEHYPDS 3180
 RESULT 16
 Y511_RICPR STANDARD: PRT; 950 AA.
 AC Q92D36;
 DT 30-MAY-2000 (rel. 39, Created)
 DT 30-MAY-2000 (rel. 39, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Hypothetical protein Rp511.

GN Rp511.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsia.
 ON NCBI_TaxID=782;
 RX STRAIN=Madrid E;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sticheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 mitochondria."
 RL Nature 396:133-140(1998).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL, AJ235272; CAA14963.1;
 DR InterPro:IPR001646; Speptide_repeat.
 DR Pfam: PF00805; Pentapeptide_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 950 AA: 108612 MW: C047F8BCF063F715 CRC64;
 Query Match 3.0%; Score 115; DB 1; Length 950;
 Best Local Similarity 18.6%; Pred. No. 2.3; Indels 268; Gaps 37;
 Matches 145; Conservative 108; Mismatches 258;
 4 EEOAKTFIDKRNHE-----AEDLFYQ--SSLASWNTNTTEENVOANNAGDKMSAFLEK 56
 120 EGGLEIKIECKRNEVYTKLDIALEVKQETDNKXTS---NILDMLTDKRFQKFN 175
 57 EOSTLAOMPIQETONTLVKIQALQALQNGSSVSEDKSKRLNTILMTSTIYSGKVCN 116
 176 EKSQKQITNH-----IKSGATELLPSDYKMSFDDIILQKPKKKLKLKFTNT---H 220
 117 PDNPOECL--LEPGLNETMANSLDYNERLWAMESRSEVQQLAPRYEEYVVLKNEMAR 174
 221 PDVKQELVANNINNPILK-KENKLEYKQELM-IFSLKEVAAQSPFLQD- 268
 175 ANHYEDYGYWRGDYEVNGVGYDYSRQQLIEDVEHTEETIKPLEYEHHAAYRAKLINAY 234
 269 --HFESY-----KIDKILDIIDPLINKIPDIKEIFDTNA----- 302
 235 PYTISPICLPAILLDMMGRFTNLYSLTVFGQKPN---IDVTDAAVDAQDQRIE 290
 303 -----PKNGVMSLEKALEWVAGDDQ--- 323
 291 KEAKFEVS--VGLPNTQGFWENS-----MLTDGNV--QKAVCHPRAM 331
 324 --LKSFPANKTILPNNALGIIENTPSVOSTNEVNFDOOMLYIVGEWMSKEIHELIA 381
 332 DLKGDRIILMCTRYT-----MDDEFLAHHEM-----GHIOYDM-AVAAQPL 373
 382 DLNKGDYMSLTGNITSIINDPSFKLADIYQSKGLPDLNLGVLEQIKNSQILKQOL 441
 374 LRNGANEGFHEAVGEINSLSAAPPKHLKSIGLSPPDFQEDNETEINFLKQALTYGILP 433
 442 INYGMENGDVYKLRKIMPIILDKRESLKV--FRDFIKGNV----- 481
 434 FTYMLEKRWAVFKGEIPKQDMKMWEMKREIVGVEPVPHDETCDPASLFHVSNDS 493
 482 -----KMKELISLTKNDPKIKRYLN-----NNRAI 507
 494 FIRYTRTYLYQFOFOALCOAAKHGBPLAKCDISNSTAGOKLFNMLRLKQSEPTVL--A 551
 508 FASIIDKTLMDLP-----GINNLDK--OELYNILPSMLNHPDELITY 547

QY 532 LENVV-----GAKN--MNVRLNTEPEFLFTMLKQDNKNSF-VGSHSTMSPYADQSIKVR 603
 Db 548 IEEERKSHYGHGAVSAIYNLAQKNYTE-----GQLPITIKAGNSGFN-YATEKVDV 599
 QY 604 ISLSALGDKAYEMNDNEMNEMTLFSSVAYAMROYFLKVNQMILFEEGEVVRANKPKPIS 662
 Db 600 FSSSRDFKDKVID-----ETIRKHLKIQN-----GKFLBGAIIILGMLSN 641
 QY 663 FNEFYAPKNVSDIIPRETEVERKAIKMSRSRIDARLNDLSLEFLGID-----PTLGP 715
 Db 642 IDFSGVSLKN-ADLFTVSTL-KDCNFKRTNLVDA-KLPDNLIMETDTYTNLDKAIPTLAP 697

RESULT 17

YAMB_SCHPO

ID YAMB_SCHPO STANDARD; PRT; 3655 AA.

AC Q10064;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C1F5.11c in chromosome 1.
 GN SPAC1F5.11c.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 NX NCBI_TaxID=4896;
 [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-972;
 RX MEDLINE-21848401; PubMed-11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby I., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovich E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens S., Sharp S.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Wellens I., Vansteelandt E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Muehlbauer S.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leclercq V., Mottier S.,
 RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallierin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez J., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrall B.G., Nurse P.,
 RA The genome sequence of Schizosaccharomyces pombe.
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: STRONG; TO YEAST YHR099W.
 CC -1- SIMILARITY: BELONGS TO THE P13/P14-KINASES FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; 268136; CA92239.1;
 DR InterPro; IPR003151; FAT.
 DR InterPro; IPR003152; FATC.
 DR InterPro; IPR000403; P13_P14_kinase.

DR InterPro; IPR001440; TPR.
 DR Pfam; PF00454; P13_P14_kinase; 1.
 DR Pfam; PF02259; FAT; 1.
 DR Pfam; PF02260; FATC; 1.
 DR SMART; SM00146; P13K; 1.
 DR PROSITE; PS50290; P13_4_KINASE_3; 1.
 DR Hypothetical protein: Transferase, Kinase.
 FT DOMAIN 3324 3655
 SO SEQUENCE 3655 AA; 420774 MW; 50475E3F3FC2124A CRC64;

Query Match 2.9%; Score 113.5; DB 1; Length 3655;
 Best Local Similarity 16.9%; Pred. No. 20;
 Matches 90; Conservative 92; Mismatches 193; Indels 157; Gaps 22;

QY 2 TIEEOAKT---FLDKRNEHAEDELFT---QSSLSAMNTNITEENYQNMNMNGDKKSAF 54
 Db 2641 TLEKATKTSLSPEFSLRRHTADALTYLNTQRMKGSVTFEFLIDEQMFSLR---RQOL 2697
 QY 55 LKE--OSTLAOMYPLOEIQNLTVKQLQALQONGSVLSSEKSKRLNTIINTWSTYSG 112
 Db 2698 PKRYQSHVSLHHQEIYEL-----QEARGLYSQINDTNININDNKL----- 2740
 QY 113 KYCNPNNQECILLPEGLNEIMANSIDYNERLMA-WESWSEYQKOLRPLEEYVLKNE 171
 Db 2741 -----RDIVVLAQGRERRLPNWVD-DIDIMEDLAWQSVFKSTINKVFLPLVSIQAO 2791
 QY 172 MARANHYEDYGDYNGDEYVNGVDGYDSRQGLDEVEHTFEIKPLYEHLAAVYAKLM 231
 Db 2792 STKNSMTNSVSYLYNGIHEL-----AWLI 2815
 QY 232 NAYPYSIPICGLPAHLIGDMGRFNTNLSLTVPEGOKPNIDVTAMVDAMDAQRIFK 291
 Db 2816 NNF-AHVAVRVHLEPVQINDL-----TKIYTL-----PNEIQEAFLEKDEQEGEYK 2862
 QY 292 EAEKFFVSGVPNNLT-----QGFWSNMLTPPGVQAVCAVCPHAPMDLCK 335
 Db 2863 SPSEWQGLDEVINTNTNLMYRNRQKAEFFLKGMEFN-RLGKRDANOAFATAVDIGS 2921
 QY 336 GDFRLMCTKTYMDLFTANHEKGIQYDMAVAAPFLLRNGAN--EGEHEVGEISLS 393
 Db 2922 G-----KAWSEWG-LYHDELFOANQDEIHNCANVSEFLQ-SSLSSS 2963
 QY 394 AATPKHLKSLGLSPDEQEDNEIFNLKQALITVGLPTFLYLEKRWVFGGEIPLKD 453
 Db 2964 NSKPLLRVLMILSVDSHSYSEV-----VSSFKSEIPLTNMIPF---IPQ- 3007
 QY 454 QMMKMKEMKRELYVEVPPEDETYCDPASLGHVSNDY-SFIRYRTTYG 504
 Db 3008 -----LLSALSHRESIHARAALILQIAKTYPOSLSLFOURTAIE 3044

RESULT 18

NEBU_HUMAN

ID NEBU_HUMAN STANDARD; PRT; 6669 AA.

AC P20929; 015346;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nebulin.
 GN NEB.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NX NCBI_TaxID=9606;
 [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE-95257391; PubMed-7739042;
 RA Labat S., Kolmerer B.;
 RT "The complete primary structure of human nebulin and its correlation
 RL to muscle structure".
 RN J. Mol. Biol. 248:308-315(1995).
 [2]

RP PARTIAL PRELIMINARY SEQUENCE.

RA MEDLINE-88284704; PubMed-3397062;
RA Zeviani M., Darras B.T., Rizzuto R., Salvati G., Betto R.,
RA Bonilla E., Miranda A.F., Du J., Samitt C., Dickson G., Walsh F.S.,
RA DiMauro S., Francke U., Schon E.A.;
RT "Cloning and expression of human nebulin cDNAs and assignment of the
RT gene to chromosome 2q31-q32.";
RL Genomics 2:249-256(1988).
RN [3]
RP STRUCTURE BY NMR OF 6610-6669.
RX MEDLINE-98179539; PubMed-9514727;
RA Pollitt A.S., Millevoi S., Gautel M., Kolmerer B., Pastore A.;
RT "SH3 in muscles: solution structure of the SH3 domain from nebulin.";
RL J. Mol. Biol. 276:189-202(1998).
CC -1- FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MAINTAINING
CC THE STRUCTURAL INTEGRITY OF SARCOMERES AND THE MEMBRANE SYSTEM
CC ASSOCIATED WITH THE MYOFIBRILS. BIND AND STABILIZE F-ACTIN.
CC -1- TISSUE SPECIFICITY: MUSCLE-SPECIFIC. LOCATED IN THE THIN FILAMENT
CC OF STRIPED MUSCLE.
CC -1- DISEASE: Defects in NEB are a cause of the autosomal recessive
CC form of nemaline myopathy (NEM).
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 178 NEBULIN REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb-sb.ch).
CC -----
DR EMBL; X83957; CAA58788.1; -;
DR EMBL; M19668; AAA59916.1; ALT_SEQ.
DR EMBL; M19669; AAA59917.1; ALT_SEQ.
DR PIR; A29979; A29979.
DR PIR; B29979; B29979.
DR PDB; IARK; 28-JAN-98.
DR PDB; INEB; 24-DEC-97.
DR Genew; HGNC:7720; NEB.
DR MIM; 161650; -;
DR MIM; 256030; -;
DR Interpro; IPR000900; Nebulin.
DR Interpro; IPR001452; SH3.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00880; Nebulin; 146.
DR PRINTS; PR00510; NEBULIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR SMART; SM00227; NEBU; 181.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00002; SH3; 1.
KW Actin-binding; Muscle protein; Cytoskeleton; Repeat; 3D-structure.
FT REPEAT 76 107 NEBULIN 1.
FT REPEAT 113 143 NEBULIN 2.
FT REPEAT 148 178 NEBULIN 3.
FT REPEAT 183 213 NEBULIN 4.
FT REPEAT 218 248 NEBULIN 5.
FT REPEAT 253 283 NEBULIN 6.
FT REPEAT 289 318 NEBULIN 7.
FT REPEAT 324 354 NEBULIN 8.
FT REPEAT 363 393 NEBULIN 9.
FT REPEAT 398 428 NEBULIN 10.
FT REPEAT 434 464 NEBULIN 11.
FT REPEAT 502 532 NEBULIN 12.
FT REPEAT 537 567 NEBULIN 13.
FT REPEAT 573 603 NEBULIN 14.
FT REPEAT 611 641 NEBULIN 15.
FT REPEAT 681 711 NEBULIN 16.
FT REPEAT 749 779 NEBULIN 17.
FT REPEAT 814 844 NEBULIN 18.
FT REPEAT 820 850 NEBULIN 19.
FT REPEAT 858 888 NEBULIN 20.
FT REPEAT 893 923 NEBULIN 21.
FT REPEAT 924 954 NEBULIN 22.

FT REPEAT 959 990 NEBULIN 23.
FT REPEAT 993 1023 NEBULIN 24.
FT REPEAT 1028 1058 NEBULIN 25.
FT REPEAT 1064 1094 NEBULIN 26.
FT REPEAT 1102 1132 NEBULIN 27.
FT REPEAT 1137 1167 NEBULIN 28.
FT REPEAT 1168 1198 NEBULIN 29.
FT REPEAT 1204 1234 NEBULIN 30.
FT REPEAT 1237 1267 NEBULIN 31.
FT REPEAT 1272 1302 NEBULIN 32.
FT REPEAT 1308 1338 NEBULIN 33.
FT REPEAT 1346 1376 NEBULIN 34.
FT REPEAT 1381 1411 NEBULIN 35.
FT REPEAT 1412 1442 NEBULIN 36.
FT REPEAT 1448 1478 NEBULIN 37.
FT REPEAT 1481 1511 NEBULIN 38.
FT REPEAT 1516 1546 NEBULIN 39.
FT REPEAT 1552 1582 NEBULIN 40.
FT REPEAT 1590 1620 NEBULIN 41.
FT REPEAT 1625 1655 NEBULIN 42.
FT REPEAT 1656 1686 NEBULIN 43.
FT REPEAT 1692 1722 NEBULIN 44.
FT REPEAT 1725 1755 NEBULIN 45.
FT REPEAT 1760 1790 NEBULIN 46.
FT REPEAT 1796 1826 NEBULIN 47.
FT REPEAT 1834 1864 NEBULIN 48.
FT REPEAT 1869 1899 NEBULIN 49.
FT REPEAT 1900 1930 NEBULIN 50.
FT REPEAT 1936 1966 NEBULIN 51.
FT REPEAT 1969 1999 NEBULIN 52.
FT REPEAT 2004 2034 NEBULIN 53.
FT REPEAT 2040 2070 NEBULIN 54.
FT REPEAT 2078 2108 NEBULIN 55.
FT REPEAT 2113 2143 NEBULIN 56.
FT REPEAT 2144 2174 NEBULIN 57.
FT REPEAT 2180 2210 NEBULIN 58.
FT REPEAT 2213 2243 NEBULIN 59.
FT REPEAT 2248 2278 NEBULIN 60.
FT REPEAT 2284 2314 NEBULIN 61.
FT REPEAT 2322 2352 NEBULIN 62.
FT REPEAT 2357 2387 NEBULIN 63.
FT REPEAT 2388 2418 NEBULIN 64.
FT REPEAT 2423 2453 NEBULIN 65.
FT REPEAT 2456 2486 NEBULIN 66.
FT REPEAT 2491 2521 NEBULIN 67.
FT REPEAT 2527 2557 NEBULIN 68.
FT REPEAT 2565 2595 NEBULIN 69.
FT REPEAT 2600 2630 NEBULIN 70.
FT REPEAT 2631 2661 NEBULIN 71.
FT REPEAT 2666 2696 NEBULIN 72.
FT REPEAT 2699 2729 NEBULIN 73.
FT REPEAT 2734 2764 NEBULIN 74.
FT REPEAT 2770 2800 NEBULIN 75.
FT REPEAT 2808 2838 NEBULIN 76.
FT REPEAT 2843 2873 NEBULIN 77.
FT REPEAT 2874 2904 NEBULIN 78.
FT REPEAT 2909 2939 NEBULIN 79.
FT REPEAT 2942 2972 NEBULIN 80.
FT REPEAT 2977 3007 NEBULIN 81.
FT REPEAT 3013 3043 NEBULIN 82.
FT REPEAT 3051 3081 NEBULIN 83.
FT REPEAT 3086 3116 NEBULIN 84.
FT REPEAT 3117 3147 NEBULIN 85.
FT REPEAT 3152 3182 NEBULIN 86.
FT REPEAT 3185 3215 NEBULIN 87.
FT REPEAT 3220 3250 NEBULIN 88.
FT REPEAT 3256 3286 NEBULIN 89.
FT REPEAT 3294 3324 NEBULIN 90.
FT REPEAT 3329 3359 NEBULIN 91.
FT REPEAT 3360 3390 NEBULIN 92.
FT REPEAT 3395 3425 NEBULIN 93.
FT REPEAT 3428 3458 NEBULIN 94.
FT REPEAT 3463 3493 NEBULIN 95.

```

FT REPEAT 3499 3529 NEBULIN 96.
FT REPEAT 3537 3567 NEBULIN 97.
FT REPEAT 3572 3602 NEBULIN 98.
FT REPEAT 3603 3633 NEBULIN 99.
FT REPEAT 3638 3668 NEBULIN 100.
FT REPEAT 3671 3701 NEBULIN 101.
FT REPEAT 3706 3736 NEBULIN 102.
FT REPEAT 3742 3772 NEBULIN 103.
FT REPEAT 3780 3810 NEBULIN 104.
FT REPEAT 3815 3845 NEBULIN 105.
FT REPEAT 3846 3876 NEBULIN 106.
FT REPEAT 3914 3944 NEBULIN 107.
FT REPEAT 3949 3979 NEBULIN 108.
FT REPEAT 4014 4044 NEBULIN 109.
FT REPEAT 4021 4052 NEBULIN 110.
FT REPEAT 4057 4087 NEBULIN 111.
FT REPEAT 4088 4118 NEBULIN 112.
FT REPEAT 4123 4153 NEBULIN 113.
FT REPEAT 4156 4186 NEBULIN 114.
FT REPEAT 4191 4220 NEBULIN 115.
FT REPEAT 4226 4256 NEBULIN 116.
FT REPEAT 4264 4294 NEBULIN 117.
FT REPEAT 4299 4329 NEBULIN 118.
FT REPEAT 4330 4360 NEBULIN 119.
FT REPEAT 4365 4395 NEBULIN 120.
FT REPEAT 4400 4430 NEBULIN 121.
FT REPEAT 4435 4465 NEBULIN 122.
FT REPEAT 4471 4501 NEBULIN 123.
FT REPEAT 4544 4574 NEBULIN 124.
FT REPEAT 4575 4605 NEBULIN 125.
FT REPEAT 4610 4640 NEBULIN 126.
FT REPEAT 4645 4675 NEBULIN 127.
FT REPEAT 4680 4710 NEBULIN 128.
FT REPEAT 4716 4746 NEBULIN 129.

```

Query Match 2.9%; Score 113; DB 1; Length 6669;
 Best Local Similarity 18.4%; Pred. No. 51;
 Matches 169; Conservative 123; Mismatches 261; Indels 364; Gaps 50;

```

50 KMSAFKEOSTLAQMTPLQEIOMLTV-KTQLQMLQONGSSVLS--DKSKRLNLTNTM 105
2387 QMWC-LPDNDVYQAKKVELESENLYKSDLEWLRGIGSPGSLAEKKNR-----A 2438
106 STYISGKVCNPNPOECILLBPGLAEIWA--NSLDYNERLW--AMESRSEYKGLARL 161
2439 SEIISSEKKYKOPDRKFTSIPDAMDIVLAKTNKRSRLYREAWDKDKTOI--HIMPD 2496
162 YEYVYVLKKNEMARANYEDYGDVWNGDYVNGVDYD-----Y 199
2497 TPPIV-----LAKANLINTSDKLYRMGYELKRRKGTDLVDAIRIKAAKASRIASEYKY 2551
200 SRG--OL-----IED-----VEHTEERKPLYEH-----LH 223
2552 KESFRKROLGHGHCARNIEDDPKMMMSMHVAKIQSDREYKKDFKWKTKSSPYDMGLVYL 2611
224 AYVRAKIMN--APRYSIPGICPAHLHGMWRFRFTNLYSLVPRPGQPNIDVTDAM-V 280
2612 AYKCOQLTVSDVCKNKLHONTCLP-----DOSVYIHA 2643
281 DQAMDAQ--RIKFEAEKFEVSVGLPMTQGFEN-----SM/LDPGANYOKAVCHP--- 328
2644 KQAVDQSDMLYKSDIQMLKGIOW--WTSGSLEDENKKNRATQTLSD--HYVHQ--HPDQF 2697
329 -----TAMDLSKGFRIIMCT-----KVMDDDL 352
2698 KFSIMDSIPMVLAKNNAITMNRHLYTEAMDCKTIVHIMDPREVLAKONKRVANSEKI 2757
353 TANHENGHIQ-----YMAVAAQPFLLRNGA-----NEGFEHVAQVIMSLSA--T 396
2758 ---YKIGLEEKRRKGYMDALPIKAKASRIASEFKYKESIRKQLGHGHCARNLRDD 2814
397 PKHLSIGLSPDOEDNETEINF-----LKOALITVGLTPFTYMLE 439

```

```

Db 2815 PKMMMSMHVAK--IQSDREYKKDFEKWKTKSSPYDMGLVLAKKCOQLTVSDVYKNNYLH 2872
QY 440 KWRM-----VFKEIIPKQMMK--WM-----EMKR--EIVG 468
Db 2873 QMTLPQOSDVYIARQAYDQSDNMYKSDL--QMRGIGWSIGSLDYEKKRATIELIS 2929
QY 469 VVEPPEDEYCDASLF-----HVSND 491
Db 2930 -----DKIYRQPPDRFFKFTSVDSLEQVLAKNNALNNKRLYTEAMDCKTQIIMPD 2982
QY 492 YSFIRY-----YTRLLYQFOQFALCOQAKHGG-----PLHKCDISNSTEAGOKL 536
Db 2983 TPETMLARONKINSEFLYKLANEE-----AKKGGYDRSDALPIVAAKASRVIISDYKX 3037
QY 537 FNMRLKSEPTLALENVGAKNNMVRPLINTEPEFLTLKDON--KNSFVGMSWSPY 595
Db 3038 -----KQGYRQLGHGHCARNIEDDPKMMMSMHVAKIQSDREYKKDFEKWKTKSSP 3089
QY 596 AD-----QSIKVRISLKSALGDKATEW-----NNEMYLPRSSVAYAMROYFLK 639
Db 3090 VDMGLVYLAKKCOQLTVSDVYKNNYL--HEWTCLPDQND-----VIHARQAYDQ 3136
QY 640 VKN-----QMLTGEEDY-----RVANLKPRISEFVYAPKNVSDIIPTEVE 683
Db 3137 SDNIYKSDLOWLRGIGFWPIGSMOYVAKCKRAAEI--LSDNIYRPP--DKLKTFSVT 3189
QY 684 KAIRMSRSRINDAFRLN 700
Db 3190 DLEQVLAK--NNALNMN 3205

```

RESULT 19

```

ID MTHR CAEEL STANDARD; PRT; 663 AA.
AC Q17693;
DE 15-JUL-1999 (Rel. 38, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable methyltetrahydrofolate reductase (EC 1.5.1.20).
GN C0668.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidea; Rhabditoidea;
OC Rhabditidae; Pelodieridae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Leinbach D.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RP REVISIONS.
RA Waterston R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADPH(+)=5,10-
CC methyltetrahydrofolate + NADPH.
CC -1- COFACTOR: FAD (by similarity).
CC -1- PATHWAY: Folate metabolism.
CC -1- SIMILARITY: BELONGS TO THE METHYLENENETETRAHYDROFOLATE REDUCTASE
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U39849; AA81048.2; -.
CC HSSP: P00394; 1B5T.
CC Wormpep: C06A8.1; C030593.
CC Interpro: IPR004621; Fadh2_euk.
CC Interpro: IPR003171; Methylrof_redctse.

```


CC -1- ALTERNATIVE PRODUCTS: 6 ISOFORMS; 1 (SHOWN HERE), 2, 3/CG-NAP, 4/YOTIAO, 5 AND 6/AAAP350; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. ISOFORM 4/YOTIAO IS HIGHLY EXPRESSED IN SKELETAL MUSCLE AND IN PANCREAS.

CC -1- DOMAIN: RII BINDING SITE. PREDICTED TO FORM AN AMPHIPHATIC HELIX. COULD PARTICIPATE IN PROTEIN-PROTEIN INTERACTIONS WITH A COMPLEMENTARY SURFACE ON THE R-SUBUNIT DIMER.

CC -1- CAUTION: REF.6 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO TWO FRAMESHIFTS IN POSITIONS 3782 AND 3811.

CC -1- CAUTION: REF.9 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FOUR FRAMESHIFTS IN POSITIONS 29, 1653, 1699 AND 1735.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AJ131693; CAB40713.1; -

DR EMBL; AB019691; BAA78718.1; -

DR EMBL; AJ010770; CA009361.1; -

DR EMBL; AF026245; AAB86384.1; -

DR EMBL; AF083037; AAD22767.1; -

DR EMBL; AC004013; AAB96867.1; -

DR EMBL; AF091711; AAD39719.1; -

DR EMBL; AB018346; BAA34523.1; -

DR EMBL; AC000066; AAC60380.1; -

DR Genew; HGNC:379; AKAP9.

DR MIM; 604001; -

DR Colled coll: Alternative splicing; Polymorphism.

FW DOMAIN 2554 2567 PKA-RII SUBUNIT BINDING DOMAIN.

FW DOMAIN 164 914

FW DOMAIN 944 1022

FW DOMAIN 1100 1185

FW DOMAIN 1253 1280

FW DOMAIN 1336 1392

FW DOMAIN 1434 1459

FW DOMAIN 1585 1659

FW DOMAIN 1857 2455

FW DOMAIN 2544 2561

FW DOMAIN 2603 2776

FW DOMAIN 3065 3092

FW DOMAIN 3124 3470

FW DOMAIN 3587 3689

FW DOMAIN 3726 3730

FW DOMAIN 203 292

FW DOMAIN 321 1010

FW DOMAIN 1846 2772

FW VARSPLIC 17 28

FW VARSPLIC 1637 1642

FW VARSPLIC 1643 3911

FW VARSPLIC 2175 2182

FW VARSPLIC 2175 2183

FW VARSPLIC 2895 2907

FW VARSPLIC 2895 2948

FW VARSPLIC 3901 3911

FW VARIANT 1347 1347

FW CONFLICT 76 76

FW CONFLICT 475 475

FW CONFLICT 554 554

FW CONFLICT 638 638

FW CONFLICT 663 663

FW CONFLICT 913 913

FW CONFLICT 956 956

FW CONFLICT 980 982

FW CONFLICT 997 997

FW CONFLICT 1001 1001

Q -> Q (IN REF. 3).

M -> I (IN REF. 3).

E -> G (IN REF. 3).

R -> S (IN REF. 3).

N -> S (IN REF. 3).

H -> N (IN REF. 3).

K -> N (IN REF. 3).

OKH -> PKP (IN REF. 1 AND 2).

Q -> P (IN REF. 1 AND 2).

Q -> P (IN REF. 1 AND 2).

FT CONFLICT 1020 1020 N -> D (IN REF. 3).

FT CONFLICT 1028 1028 V -> E (IN REF. 3).

FT CONFLICT 1626 1626 R -> P (IN REF. 1 AND 2).

FT CONFLICT 1703 1703 N -> T (IN REF. 3).

FT CONFLICT 1707 1707 V -> G (IN REF. 3).

FT CONFLICT 1802 1803 MISSING (IN REF. 5).

FT CONFLICT 1843 1843 A -> P (IN REF. 3).

FT CONFLICT 1956 1956 I -> V (IN REF. 3).

FT CONFLICT 2027 2027 V -> D (IN REF. 5).

FT CONFLICT 2157 2158 EI -> HE (IN REF. 7).

FT CONFLICT 2169 2169 E -> V (IN REF. 3).

FT CONFLICT 2514 2514 L -> R (IN REF. 3).

FT CONFLICT 2851 2851 I -> N (IN REF. 8).

FT CONFLICT 2957 2957 E -> D (IN REF. 3).

FT CONFLICT 2983 2983 P -> S (IN REF. 3, 7 AND 8).

FT CONFLICT 3087 3087 Q -> H (IN REF. 3).

FT CONFLICT 3218 3218 Q -> H (IN REF. 3).

FT CONFLICT 3307 3309 BSE -> OSO (IN REF. 3).

FT CONFLICT 3751 3751 P -> A (IN REF. 3).

FT CONFLICT 3833 3833 T -> S (IN REF. 3).

SQ SEQUENCE 3911 AA; 453664 MM; 3FB1CB1C819B47AA CRC64;

Query Match 2.9% Score 112; DB 1; Length 3911;

Best Local Similarity 20.5% Pred. No. 28; Mismatches 122; Indels 84; Gaps 15;

Matches 68; Conservative 57; Mismatches 122; Indels 84; Gaps 15;

QY 11 LDKFHEADLFYQSS---LASWYNTNTEENYNNAGDKWSAFLEKOSTLAQWYP 66

DB 1111 LNVKSEQNDLRLQWEAQRICSLVYSTHVDQVR-EYENKDKALCSLKEELFAQBEK 1169

QY 67 LOEIQNL-TVKIQLOALQONG-----SSVISEDKSKRLNTILTMSTIYSTG 112

DB 1170 IRELQKHOLEQTKTQETGEGKPLHLIGLKQKAVSECSYFLQTLCSVLGEYTPA 1229

QY 113 KYC-----NPDRPOCLL--LEPGNEIMANSIDYNE----- 142

DB 1230 LKCEVNAEDKENSQDIYSNEDPELODTRYEVQDQEMHTLLNKVTEYKLLVYLQTRL 1289

QY 143 -RIMAMES--WRSEVQKQLRPLYE-EYVLKEMARAHYEDYGDYRGDYEVNGVDGYD 198

DB 1290 SKIWGQDTGKMLKEGEEHLPRETEFLSHSQMNL-----DIDVN-----H 1333

QY 199 YRGQILIEVEIT-FEIKPLYLEHAYRAKLMAYSYISPIGCLPAHLIGDMGRFW 257

DB 1334 KSKLSLDPLETKLEBQVLESLSISLQOOLKETEONTEAFELHCLQKRLQA----- 1386

QY 258 TNLISLTVFGQKPNIDYTDANVQDAMDOR 288

DB 1387 --VSESTVP-----PSLPVDSDVITTES-DAQR 1410

Search completed: May 26, 2003, 17:54:18

Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 16:58:14 ; Search time 69 Seconds

(without alignments)
2150.057 Million cell updates/sec

Title: us-09-978-385-2_COPY_19_738

Perfect score: 3869
Sequence: 1 STIEBOAKTFIDKFNHEAD.....DNSLEFLGIQPTLGPPNPP 720

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertbrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	804	4	09UPZ6
2	3869	100.0	805	4	09NRA7
3	3866	99.9	805	4	09NRA7
4	3251	84.0	798	11	09NRA7
5	3251	84.0	805	11	09NRA7
6	1461	37.8	353	11	09NRA7
7	1329	34.3	732	6	09GLN7
8	1329	34.3	1304	6	09GLN7
9	1310	33.9	1313	11	09EOM9
10	1303	33.7	694	4	015540
11	1087	28.1	615	5	09NDS8
12	1069	27.2	648	5	09NDS8
13	1054	26.6	660	5	017248
14	1027.5	26.5	630	5	024222
15	1025.5	26.5	630	5	09VLJ6
16	930	24.0	844	5	08SXX2

17	841	21.7	792	5	09VT1
18	762	19.7	249	11	09D836
19	638.5	16.5	907	5	018581
20	491	12.7	661	5	09V520
21	458	11.8	611	5	09V1V2
22	434.5	11.2	202	11	064603
23	407.5	10.5	628	5	09S079
24	407.5	10.5	628	5	09W021
25	306.5	7.9	121	6	09BDG1
26	280.5	7.2	135	4	016425
27	272	7.0	222	4	09HBJ8
28	271	7.0	222	11	09S633
29	265	6.8	222	11	09S634
30	251.5	6.5	157	11	061265
31	239	6.2	75	5	09TK66
32	157	4.1	532	16	09KEV0
33	154	4.0	502	16	08Y616
34	152.5	3.9	734	5	09NKE3
35	147	3.8	54	6	09S161
36	147	3.8	502	16	09ZAC3
37	139.5	3.6	987	16	08RW66
38	139	3.6	608	16	09PR80
39	136	3.5	611	16	09PPW8
40	130.5	3.4	461	2	09K152
41	125.5	3.2	600	16	08R670
42	125	3.2	3571	10	09SL27
43	125	3.2	3571	10	09SL27
44	124.5	3.2	987	11	061636
45	124.5	3.2	3429	11	008614

ALIGNMENTS

RESULT 1	09UPZ6	PRELIMINARY;	PRT;	804 AA.
ID	09UPZ6	01-MAY-2000 (TREMBLrel. 13, Created)		
AC	09UPZ6	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAY-2000	(TREMBLrel. 19, Last annotation update)		
DE	Hypothetical 92.3 kDa protein (Fragment).			
GN	DKEZF434A014.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=TESTIS;			
RA	Wamunt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;			
RL	Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AL10224; CAB53682.1; -			
DR	MEROPS: M02.006; -			
DR	InterPro: IPR001548; Peptidase_M2.			
DR	InterPro: IPR000130; Zn_MTPeptide.			
DR	Pfam: PF01401; Peptidase_M2; 1.			
DR	PRINTS: PR00791; PEPDIPASPA.			
DR	ProDom: PD004184; Peptidase_M2; 1.			
DR	PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.			
KW	Hypothetical protein.			
FT	NON_TER			
SQ	SEQUENCE	804 AA; 92340 MW; 91FF391074CB5DA9 CRC64;		
Query Match	100.0%; Score 3869; DB 4; Length 804;			
Best Local Similarity	100.0%; Pred. No. 6.1e-275;			
Matches	720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 STIEBOAKTFIDKFNHEADLFYOSLSAAMVNTNTEENVONNANNGDKWAFKRGST 60			
DB	18 STIEBOAKTFIDKFNHEADLFYOSLSAAMVNTNTEENVONNANNGDKWAFKRGST 77			
OY	61 LAQWVPIQEIQNLVVKIQQLALQNGSSVLSSEKSKLNTLITLMTSTTSTGKVCNPDP 120			

```

Db 78 LAQWPIQEIQNLTKVQLQALQNGSSVLSSEKSKRLNTILNTSTIYTGKVCNDNP 137
QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVSKOLRPLYEEVYLKNEARAHNHED 180
Db 138 QECILLEPGLNEIMANSIDYNERLWAMESRSEVSKOLRPLYEEVYLKNEARAHNHED 197
QY 181 YGDYWRGDEYVNGVGYDYSRQGLIEDVEHTFEETIKPLYEHLHAYRAKLMAAPSYISP 240
Db 198 YGDYWRGDEYVNGVGYDYSRQGLIEDVEHTFEETIKPLYEHLHAYRAKLMAAPSYISP 257
QY 241 IGCLEPAHLIGDMGKGFNTNLSLYVPGQKPNIDVTAMDADQAPRIFKAKEFPVS 300
Db 258 IGCLEPAHLIGDMGKGFNTNLSLYVPGQKPNIDVTAMDADQAPRIFKAKEFPVS 317
QY 301 GLPNMTQGFWEKSMILDPGNVOKAVCHPTAMDLGKGFRIIMCTKYTMDFLTAHHEMCH 360
Db 318 GLPNMTQGFWEKSMILDPGNVOKAVCHPTAMDLGKGFRIIMCTKYTMDFLTAHHEMCH 377
QY 361 IQYDMAVAAQPLRLNGANGEGHEAVGEIMSLSAATPKHLKSGILSPDQEDNETEINF 420
Db 378 IQYDMAVAAQPLRLNGANGEGHEAVGEIMSLSAATPKHLKSGILSPDQEDNETEINF 437
QY 421 LKQALTYIGTLPTFTMLEKRMWVKGELPKDQMKKMKMKREIYGVVEVPHDETYC 480
Db 438 LKQALTYIGTLPTFTMLEKRMWVKGELPKDQMKKMKMKREIYGVVEVPHDETYC 497
QY 481 DPASLFHVSNDYSFIRYTRTYLQFOFQALCOAAKHGSPILKCDISNSTEAGOKLFNML 540
Db 498 DPASLFHVSNDYSFIRYTRTYLQFOFQALCOAAKHGSPILKCDISNSTEAGOKLFNML 557
QY 541 RLGKSEPTLALENVGAKNMNVRPLNTEPELFTWLDKQNKNSVGVSTMSPYADDSI 600
Db 558 RLGKSEPTLALENVGAKNMNVRPLNTEPELFTWLDKQNKNSVGVSTMSPYADDSI 617
QY 601 KYRISLSKALGDKAYEMNDENKYLFRSSVAYAMROYFLKYNOMILFGEEDVYVANKR 660
Db 618 KYRISLSKALGDKAYEMNDENKYLFRSSVAYAMROYFLKYNOMILFGEEDVYVANKR 677
QY 661 ISNFEVTPAKNVSDIIPTEVEKALRMSRSRINDAFRLNDSLEFLGIQPLGPNO 720
Db 678 ISNFEVTPAKNVSDIIPTEVEKALRMSRSRINDAFRLNDSLEFLGIQPLGPNO 737

RESULT 2
Q9NRA7 PRELIMINARY; PRT; 805 AA.
ID 09NRA7:
AC 09NRA7:
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE Angiotensin converting enzyme-like protein (ACE2)-related
DE carboxypeptidase (ACE2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RC TISSUE-LYMPHOMA;
RA Turner A.J., Hooper N.M., Hyde R.J., Christie G., Karian E.,
RA Turner A.J.,
RA "A Human Homolog of Angiotensin Converting Enzyme - Cloning and
RA Functional Expression As A Captopril-Insensitive Carboxypeptidase.";
RA Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Donoghue M., Hsieh F., Baronas F., Godbout K., Gosselin M.,
RA Stagliano N., Donovan M., Woolf B., Robison K., Jayaseelan R.,
RA Breitbart R.E., Acton S.;
RA "A novel ACE-related carboxypeptidase (ACE2) converts angiotensin I to
RA angiotensin-1-9.";
RL Circ. Res. 0:0-0(2000).

```

```

DR EMBL; AF241254; AAF78220.1; -
DR EMBL; AF291820; AAF99721.1; -
DR MEROPS; M02.006; -
DR InterPro; IPR001548; Peptidase_M2.
DR InterPro; IPR000130; Zn_peptidase.
DR Pfam; PF01401; Peptidase_M2; 1.
DR PRINTS; PR00791; PEPTIDASEA.
DR ProDom; PD004184; Peptidase_M2; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Carboxypeptidase.
SQ SEQUENCE 805 AA; 92462 MW; 8EE6EB0A931550E8 CRC64;

Query Match 100.0%; Score 3869; DB 4; Length 805;
Best Local Similarity 100.0%; Pred. No. 6,1e-275;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIEQATFELDKRNEHEADLFYQSLASWNTNTTEENVQNNMNGKMSAFLEKQST 60
Db 19 STIEQATFELDKRNEHEADLFYQSLASWNTNTTEENVQNNMNGKMSAFLEKQST 78
QY 61 LAQWPIQEIQNLTKVQLQALQNGSSVLSSEKSKRLNTILNTSTIYTGKVCNDNP 120
Db 79 LAQWPIQEIQNLTKVQLQALQNGSSVLSSEKSKRLNTILNTSTIYTGKVCNDNP 138
QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVSKOLRPLYEEVYLKNEARAHNHED 180
Db 139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVSKOLRPLYEEVYLKNEARAHNHED 198
QY 181 YGDYWRGDEYVNGVGYDYSRQGLIEDVEHTFEETIKPLYEHLHAYRAKLMAAPSYISP 240
Db 199 YGDYWRGDEYVNGVGYDYSRQGLIEDVEHTFEETIKPLYEHLHAYRAKLMAAPSYISP 258
QY 241 IGCLEPAHLIGDMGKGFNTNLSLYVPGQKPNIDVTAMDADQAPRIFKAKEFPVS 300
Db 259 IGCLEPAHLIGDMGKGFNTNLSLYVPGQKPNIDVTAMDADQAPRIFKAKEFPVS 318
QY 301 GLPNMTQGFWEKSMILDPGNVOKAVCHPTAMDLGKGFRIIMCTKYTMDFLTAHHEMCH 360
Db 319 GLPNMTQGFWEKSMILDPGNVOKAVCHPTAMDLGKGFRIIMCTKYTMDFLTAHHEMCH 378
QY 361 IQYDMAVAAQPLRLNGANGEGHEAVGEIMSLSAATPKHLKSGILSPDQEDNETEINF 420
Db 379 IQYDMAVAAQPLRLNGANGEGHEAVGEIMSLSAATPKHLKSGILSPDQEDNETEINF 438
QY 421 LKQALTYIGTLPTFTMLEKRMWVKGELPKDQMKKMKMKREIYGVVEVPHDETYC 480
Db 439 LKQALTYIGTLPTFTMLEKRMWVKGELPKDQMKKMKMKREIYGVVEVPHDETYC 498
QY 481 DPASLFHVSNDYSFIRYTRTYLQFOFQALCOAAKHGSPILKCDISNSTEAGOKLFNML 540
Db 499 DPASLFHVSNDYSFIRYTRTYLQFOFQALCOAAKHGSPILKCDISNSTEAGOKLFNML 558
QY 541 RLGKSEPTLALENVGAKNMNVRPLNTEPELFTWLDKQNKNSVGVSTMSPYADDSI 600
Db 559 RLGKSEPTLALENVGAKNMNVRPLNTEPELFTWLDKQNKNSVGVSTMSPYADDSI 618
QY 601 KYRISLSKALGDKAYEMNDENKYLFRSSVAYAMROYFLKYNOMILFGEEDVYVANKR 660
Db 619 KYRISLSKALGDKAYEMNDENKYLFRSSVAYAMROYFLKYNOMILFGEEDVYVANKR 678
QY 661 ISNFEVTPAKNVSDIIPTEVEKALRMSRSRINDAFRLNDSLEFLGIQPLGPNO 720
Db 679 ISNFEVTPAKNVSDIIPTEVEKALRMSRSRINDAFRLNDSLEFLGIQPLGPNO 738

RESULT 3
Q9BYF1 PRELIMINARY; PRT; 805 AA.
ID 09BYF1:
AC 09BYF1:
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE Angiotensin converting enzyme-like protein (ACE2)-related
DE carboxypeptidase (ACE2).

```

[illegible]

THEORY OF THE FIRM

11

Db 439 LKQALITVGLPTFTYMLEKRWAVFGEIPEKQOMKMKWEMKREIVGVPELPHDETYC 498
 OY 481 DPASLEFHSNDYSFTRYYTRITLYOFQFOEALCOAKHEGPIHKCDISNSTEAGOKLFNML 540
 Db 499 DPASLEFHSNDYSFTRYYTRITLYOFQFOEALCOAKHEGPIHKCDISNSTEAGOKLFNML 558
 OY 541 RLKSEPTLALENVYGAKNMVRPLNYEPELFTWLDKONKSFVGNSTDSWSPYADQSI 600
 Db 559 SIGNSEPWTKALENVYGAKNMVRPLNYEPELFTWLDKONKSFVGNSTDSWSPYADQSI 618
 OY 601 KVRISLKSALGDKAYEWNDEMYLFRSSVAYAMROYFLKYNOMLFEEDVRYANLKPR 660
 Db 619 KVRISLKSALGANAAYEWNNEMFLFRSSVAYAMRKYSIINKQVPLEEDVRYANLKPR 678
 OY 661 ISPNFVTPAKNVSDIIPRETEVEKAIKMSRSRINDAFRLNDSLEFLGIQPTLPPNPP 720
 Db 679 VSPFFVTSPQNVSDIIPRETEVEKAIKMSRGRINDVFLGNDNSLEFLGIHPTLEPPYQPP 738

RESULT 5

08R010 PRELIMINARY; PRT; 805 AA.
 ID 08R010: 01-JUN-2002 (TREMBlrel. 21, Created)
 AC 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE RIKEN CDNA 2010305L05 gene.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC026801; AAH26801.1;
 SQ SEQUENCE 805 AA; 92367 MW; DB883AAC966A8D9 CRC64;

Query Match

Best Local Similarity 84.0%; Score 3251; DB 11; Length 805;
 Matches 598; Conservative 49; Mismatches 73; Indels 0; Gaps 0;

1 STIEBOAKFLDKFHEAEDLFYQSSLASWNTNTITEENQANNNGDKSAFLKQST 60
 Db 19 SLTEENAKTFLNNFNOEAEDLSYQSSLASWNTNTITEENQANNNGDKSAFLKQST 78
 OY 61 LAQMTPLQETQNTLVKQLQALQOQSSSVLSDESKRLNTLNTSTYSGKVCNPNP 120
 Db 79 TAQSFSLQETQPTTIKROLQALQOQSSSVLSADKKKQNTLNTSTYSGKVCNPNP 138
 OY 121 QECILLEPGINEIMANSIDYNERLWAMESWSEVQKQPLYEYEVVLKNNMAYANHYED 180
 Db 139 QECILLEPGIDEIMASTDYNSRLWAMEGMALEVQKQPLYEYEVVLKNNMAYANHYED 198
 OY 181 YGDYWRGDYEVNGVDGYDYSRGQOLIEDVHTEFEIKPLYEHLHAYVRRAKNNAPSYISP 240
 Db 199 YGDYWRGDYEAEGADGYNNRNQOLIEDVHTEFEIKPLYEHLHAYVRRAKNNAPSYISP 258
 OY 241 IGGPLPAHLGDMGGRFTNLVSLTFVFGQKPNIDVTADAVDQAMDQRIFFKAKEFYVS 300
 Db 259 TGGPLPAHLGDMGGRFTNLVSLTFVFGQKPNIDVTADAVDQAMDQRIFFKAKEFYVS 318
 OY 301 GLPMTQGFWENSMTLDPGNVQKAVCHPRAWDLGKDRITLACTVYTMDDFTLHHEKGH 360
 Db 319 GLPMTQGFWANSMTLEPADGRKVCVHPRAWDLGKDRITLACTVYTMDDFTLHHEKGH 378
 OY 361 IQYMAVYAAQPELIRNANNGEHEAVGEISLNAATPKHLKISGLISPDQDNTEIINF 420
 Db 379 IQYMAVYAAQPELIRNANNGEHEAVGEISLNAATPKHLKISGLISPDQDNTEIINF 438
 OY 421 LKQALITVGLPTFTYMLEKRWAVFGEIPEKQOMKMKWEMKREIVGVPELPHDETYC 480

RESULT 6

099N70 PRELIMINARY; PRT; 353 AA.
 ID 099N70: 01-JUN-2001 (TREMBlrel. 17, Created)
 AC 099N70: 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Angiotensin-converting enzyme-related carboxypeptidase.
 GN ACE2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Komatsu T., Sugano S., Imai J., Suzuki Y., Hanaoka K., Ymada Y.,
 RA Hide M., Tanigami A., Muroi S.;
 RT "Molecular cloning, mRNA expression, and chromosomal localization of
 RT mouse Angiotensin-converting Enzyme-Related Carboxypeptidase."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

Query Match

Best Local Similarity 37.8%; Score 1461; DB 11; Length 353;
 Matches 268; Conservative 22; Mismatches 43; Indels 0; Gaps 0;

1 STIEBOAKFLDKFHEAEDLFYQSSLASWNTNTITEENQANNNGDKSAFLKQST 60
 Db 19 SLTEENAKTFLNNFNOEAEDLSYQSSLASWNTNTITEENQANNNGDKSAFLKQST 78
 OY 61 LAQMTPLQETQNTLVKQLQALQOQSSSVLSDESKRLNTLNTSTYSGKVCNPNP 120
 Db 79 TAQSFSLQETQPTTIKROLQALQOQSSSVLSADKKKQNTLNTSTYSGKVCNPNP 138
 OY 121 QECILLEPGINEIMANSIDYNERLWAMESWSEVQKQPLYEYEVVLKNNMAYANHYED 180
 Db 139 QECILLEPGIDEIMASTDYNSRLWAMEGMALEVQKQPLYEYEVVLKNNMAYANHYED 198
 OY 181 YGDYWRGDYEVNGVDGYDYSRGQOLIEDVHTEFEIKPLYEHLHAYVRRAKNNAPSYISP 240
 Db 199 YGDYWRGDYEAEGADGYNNRNQOLIEDVHTEFEIKPLYEHLHAYVRRAKNNAPSYISP 258

DE Ecdysteroid-inducible angiotensin-converting enzyme-related
DE product.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Dityrsta;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN=C108; TISSUE=WING IMAGINAL DISK;
RX MEDLINE=20556050; PubMed=11102839;
RA Quan G.-X., Mita K., Okano K., Shimada T., Uga Jin N., Xia Z., Goto N.,
RA Kanke E., Kawasaki H.;
RT "Isolation and expression of the ecdysteroid-inducible angiotensin-
RT converting enzyme-related gene in wing discs of Bombyx mori.";
RL Insect Biochem Mol Biol. 91:97-103(2001).
DR EMBL: AB026110; BAA97657.1; -.
DR MEROPS: M02.002; -.
DR InterPro: IPR001548; Peptidase_M2

DR Pfam: PF01401; peptidase_M2; 1.
DR PRINTS: PR00791; PEPDPTASEA.
DR PRODOM: PD004184; peptidase_M2; 1

DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 648 AA; 74917 MW; 9A740AA9FCACEBF0 CRC64;

Query Match	28.1%;	Score 1087;	DB 5;	Length 648;
Best Local Similarity	36.1%;	Pred. No. 3.5e-71;		
Matches 228;	Conservative 122;	Mismatches 740;		

QY

11 LDKFNHEADLFY----- : : : : . . : : : . . :	SSLASWNTNTNTEENVOMNNA GDKWSAFLKEQS 59
---	--

D6	29	LEAREHEAREYMLHDKATGLKRNRRASLAWEYTSNITKEENEKSICQTHLELSRQEKAAW	88
OY	60	TLAQMTPLOEIGNLVKVLQALQOONSSVLSDDSKRRLNTIITNMGSTVSGGVCVNDY	110

[illegible]

Dd

```
120 QEC-LLEPBGNEIMANSLDYNERLMAWESRSEVGKQRLPYEEYVVLKNEMARANYH 178  
   : | | | : | : | | | : : : : : | | | | :  
149 ESKDISLEPETTEIFSTSDPEELKHAVEHHNAGATAKNFTDYVLNLNEAAKNGF 208
```

QY 179 EDYGDYWRGDIYEVNGVDGYDYSRGQLIEDVEHTFEETIKPLXHLHAYVRATIMNAY-PSY 237
:: ::|::| : | :||::||| :| :|
Dd 209 DNVAENQSEFEVDPDEE----- : :||||| :||| :| :|

QY

238 ISPICGLPAHLGDMGRFWNLYSITVPGQKPNIDVTDAVDQAAMDARIKEAEKF 297
| | : ||||| : : : : :
-----EQLAKLMEDVKPLVQQHAYVRKKLRDKIGDKV 258
.....GGCGLLEDE

Db 259 VSARGEPIPAHLGNMWAQTNNNIESEFTRPIPDKKETIDVQAMRDQNTYTPMKMFQMSDEFF 318

Db

319 RSLNLTAMPEKFKNSIIEKPTD-REIVYCHASAMWDFDGEDFRINQCTIVDYEIFQTTH 377

378 EMGIIQYIQTROPVVERDANGNGFHEAVGDITALLSYSPKHILRRVIGATGD--ADPDET 426

[illegible]

477 ETCYDPA^{SL}LFHVSNDYSFIRYYRT^{LT}LQFQFQALCQ-AAKHG-----PLHKDISNST 530

497 EDDFDAAKYHVSNNVEIARYIYSFIQFQHRGVCQLAGEHAAAGDPNKKLVCCDIYSV 556

557 AAGNALNMLKSSSKPDPDAHALTGCREKADGLLEYFRPHLDWLAENGRTEGHIGW 616

```

589 STDWSPYADQSIKVRISLSKALGDKAYEMNDN 620
      | | ::: | : : |
617 EPTNMEYCTP SOLSELNVEKEPSSSPATQOSDS 648

```

RESULT 12

09NKE4 PRELIMINARY; PRF: 615 AA.
 ID 09NKE4; 09YV3; Created)
 DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
 DE Ance protein (LJ1258P).
 NC ANCE OR CG8827.
 GN Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE-99403001; PubMed-10471707;
 RA Ashburner M., Mista S., Roote J., Lewis S.E., Blazee R., Davis T.,
 Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
 Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
 Celinker S., Rubin G.M.;
 RA "An exploration of the sequence of a 2.9-Mb region of the genome of
 Drosophila melanogaster: the Adh region.";
 RT Genetics 153:1179-219(1999).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE-99403001; PubMed-10471707;
 RA Ashburner M., Mista S., Roote J., Lewis S.E., Blazee R., Davis T.,
 Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
 Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
 Celinker S., Rubin G.M.;
 RA "An exploration of the sequence of a 2.9-Mb region of the genome of
 Drosophila melanogaster: the Adh region.";
 RT Genetics 153:1179-219(1999).
 RL [3]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-20196006; PubMed-10731132;
 RX Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.H.C., Blazee R.G., Champe M., Pfeiffer B.D.,
 Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
 Beeson K.Y., Benos P.V., Bernier J., Brockstein P., Brotler P.,
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 Deodson B., Delcher A., Deng Z., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
 Foadi C., Gabrielian A.E., Garg N.S., Gehlert W.M., Glasser K.,
 Glodier A., Gong F., Gotzelli J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Helman T.J., Hewlitt J., Hickey J., Hinkley J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jaffe M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhong L.,
 Zhong X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of Drosophila melanogaster";
 RT Science 287:2185-2195(2000).
 RL [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Stapleton M., Brockstein P., Hong L., Aghayani A., Carlson J.,
 Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 Raunoo J., Pacle J., Parasas V., Park S., Phouenavong S., Wan K.,
 Yu C., Lewis S.E., Rubin G.M., Celinker S.;
 RA Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.
 RL EMBL: AE003408; AAF44834.1; -;
 DR EMBL: AE003641; AAF53353.2; -;
 DR EMBL: AY061129; AAL28677.1; -;
 DR FlyBase: FBgn0012037; Ance.
 DR InterPro: IPR001548; Peptidease_M2.
 DR InterPro: IPR001130; Zn_MTPeptidease.
 DR Pfam: PF01401; Peptidease_M2; 1.
 DR PRINTS: PR00791; PEPDIPASEA.
 DR ProDom: PD004184; Peptidease_M2; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKOWN.1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKOWN.1.
 DR SEQUENCE 615 AA; 70913 MW; 9E3691BC51D6C48 CRC64;
 Query Match 27.64; Score 1069; DB 5; Length 615;
 Best Local Similarity 36.38; Pred. No. 6; 8e-70;
 Matches 217; Conservative 120; Mismatches 239; Indels 22; Gaps 9;
 4 EQAKTFLDFRNEAEEDLFYQSSLSKNTNTNTNTEENONMNRAGKNSAPFKKEOSTLAQ 63
 22 EQAKTELENKELAKRTVETEAAYAGSNTDENKKETISAELEKFEVASTDT 81
 64 MYPLQELONLTKLOLALOONGSVSEDSKRLNTIINTWSTIYSTGVCNPNPDR 123
 82 KFWKSTYQSDKQKFAKRLTKYALPREDYAEILDITLSAMSNPAKVKCYCDKSTWC 141
 124 -LLEPGELNEMANSIDYNERLWAMESRSEVQKOLRPLYEEVYVKNENARHNEDG 182
 142 DLALPELEEVYSRDEELAYRREFFYKAGVAGSQSERVEYNTAKAKNTTSSA 201
 183 DYWRDDEVNGVDVGYDSRGOLIDVETFEETKPLETHLHAYVRAKLNAV-PSYISPT 241
 202 EAWDEYE-----DITFEQLEDI---FADIRPLYQIGHYRFRKHYGDVAVSET 251
 242 GCLPAHLIGDMGRWIMYLSLVPFGQPNIDVTAMDQADQADQIFKEAKFEVSG 301
 252 GFLPHILGNMAQOMSEIADIVSPPEKPLVDVSAEKEKGYPPLMFQMGDPFTSMN 311
 302 LPMVQGFENMLDPCGVQKACHPTAMDG-KGDFRLMCTKVTMDFLAHHEMGH 360
 312 LKRLPDQFDEKIIETKPTDGRDLVCHASAMDFLYTDVRIQCIRVQDQFLVHRLGH 371
 361 IQYDMVYAAPFLRNGANGNEFHEAVGEIMTSIAATPKHKLSTGLSPDQEDNETINF 420
 372 IQYFLOYQHPFYRRGANGFHEAVGDLSTLSTVPHKRIIGLKL-DYVRDEARINQ 430
 421 LKQALTYGTLPLFTYMLEKRWMPKRGIPRDOMKRWMEKRIYGVVPHDETYC 480
 431 LFTLADKLVFLPFAFTMKYRNSLFRGVNDANNCAPFKLRDYSIEPPVRSKEDF 490
 481 DPASLTVSNDYFIRYTRTYQFQFQALC-QAKH-----EGPLHKDISNTEGQ 534
 491 DAPAKYIISDVETIKIVSFIIOFQYKACIKAGQYDPDVAEPLDNCIDYGSAAAGA 550
 535 KLFNMLRLGSEPTLALENVVAKAMNRPPLANTPEPFLTWLKOON--KNSFVMS 590
 551 AFHMLSMGASKPMPDLEAFNGERIMSGKALAEYEPRLVWLEAKNKNNYHIGWT 608

RESULT 13
ID 017248 PRELIMINARY; PRT; 660 AA.
AC 017248;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE Angiotensin-converting enzyme-like protein precursor.
GN Boophilus microplus (cattle tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Boophilus.
OX NCBI_TaxID=6941;
RN [1]
RP SEQUENCE FROM N.A.
RA Whitfield P.L., Johnson M.C., Smith D.R.J., Snelson V.J., Clifton G.F.,
RA Brown G.S., Cairns D., Foy A.B., Irving D.O.,
RT "The Cloning of an Angiotensin-Converting Enzyme cDNA from the Cattle
RT Tick Boophilus microplus."
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U62809; AAB04998.1; -
DR InterPro: IPR001548; Peptidase_M2.
DR Pfam: PF01401; Peptidase_M2; 1.
DR PRINTS: PR00791; Peptidase_M2.
DR PRODOM: PD004184; Peptidase_M2; 1.
KW signal.
FT SIGNAL. 1 29 POTENTIAL.
FT CHAIN 30 660 POTENTIAL.
SQ SEQUENCE 660 AA; 75257 MW; 6F164CF70C938B63 CRC64;
Query Match 27.2%; Score 1054; DB 5; Length 660;
Best Local Similarity 37.7%; Pred. No. 9, 5e-69;
Matches 224; Conservative 104; Mismatches 234; Indels 32; Gaps 12;
QY 4 EDOAKTFIDKFNHEADLFYQSSLSAWNTNTEENVANNMN-AGDKMSAFKQSTLA 62
DB 49 EAMGVAFIEGLNDPTTINNVSSSMDASNTIYN-QMKNKSTEVSKRERPGITA 107
QY 63 QMYPLQELONLVKQLQALQNGSSVSEDSKRLNTLTMTSTYTGKVC--NPDN 119
DB 108 KRFDMHNFNDLSLKRFRVATIGLALPDDELNATSLSSKMAIYGSTKVTYVKDML 167
QY 120 POECLLEPGLNEIANSIDYERLWAMESRSEVGKOLRLPYEEYVLKNGMAANYE 179
DB 168 F-----LEPDLTRNNKEVGNVYKQLQTLAMNANVGPRIKQYIYILSNFAASLDYD 222
QY 180 DYGDYWRGDEYVNGVDYDSRGQLEDEYHTEFEIKLYEHLNAYRAKLNAPSYIS 239
DB 223 NIKSAMLSDETE-----NMTETVDKLMEDLSPLYKKLHAYMKLREIYGRLP 272
QY 240 PIGCDPAHLIDGMGRFTNLY-SLTVPFGOKPNIDVDANVQAMQRIKFEAKETV 298
DB 273 EGGTTPAHLIGNMMAQEWGTLPRLTME--DKP-LDISKTVAEQKWDQKMFHAEDFT 329
QY 299 SVGLPNNMTQGEWNSMLTDPGVQKAVCHPTAMDLSGK-DEPILMCTKVTYDDFLTAHE 357
DB 330 SIGLDNMTSEFMSKSIITKRPD-RIQGHASAMNNYNDDRIRKICTPSPVEELTYVHE 388
QY 358 MGHIOYDMAAQAQAPLLNNGANEGHEVGEIMISAAFPRLHLSIGLSLSPDEQNETE 417
DB 369 MGHETIYQVYKHLNHLQEGANEGHEAVGDLIALSVATKTHYGLSLKFP---TDKYN 445
QY 418 INFLKQALITVGLTFEYMLKEWRMAYFEKGEIPRDOMKMKWMEKKREIVAVPEYPRDE 477
DB 446 VDLILMSLDKIALFLPRGILDKRWITFTGETPDKNENKFWERIKYQVSPPRKNE 505
QY 478 TYCPBASLFEVNSYSFIRYTRLYQFOFOALCOAK--HEGPIHKCDISNSTAGQ 534
DB 506 SEFGGAGYHVALHVPYLRKYFAVILQFOFHEHCLTVAKKVDEHHPFECDIYEKNAGD 565
QY 535 KLFNMLRLGKSEPTLALENVVGAKNMNVRLNVEFLFTWLKDQKNSEVGM 588

DB 566 VLKGLSLGRSKPWPVLEIMAGTROMSASSLKTIYEPLEKWLDERIKNEYVGM 619
RESULT 14
ID 024222 PRELIMINARY; PRT; 630 AA.
AC 024222;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE METALLOPEPTIDASE (Angiotensin-converting enzyme-related protein).
GN ACER OR CG10593.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-97128790; PubMed=8973330;
RA Taylor C.A.M., Coates D., Shirras A.D.;
RT "The Acer gene of Drosophila codes for an angiotensin-converting
RT enzyme homologue."
RL Gene 181:191-197(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BERRELEY;
RA Celniker S.E., George R.A., Galle R., Sylrakas R.R., Hoskins R.A.,
RA Agbayani A., Arcina T.T., Baxter E., Blazey R.G., Chavez C., Chev M.,
RA Doyle C.M., Farfan D.E., Flanagan J., Houston K.A., Humast S.R.,
RA Karra K., Kearney L., Kim S.H., Lee B., Lomoton M.A., Mak J.,
RA Marz P., Mok M.S., Moshrefi A.R., Moshrefi M., Nixon K., Pacled J.M.,
RA Park S., Pfeiffer B., Pouch E., Snir E., Thomey B., Wan K.H.,
RA Whitelaw K.R., Yee A., Zhang R., Zieren L., Kimmel B.E.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: X96913; CAA6532.1; -
DR FLYBase: FBgn0016122; Acer.
DR InterPro: IPR001548; Peptidase_M2.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01401; Peptidase_M2; 1.
DR PRINTS: PR00791; Peptidase_M2.
DR PRODOM: PD004184; Peptidase_M2; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
FT PROSITE 236 236 L->F (IN REF. 2).
FT CONFLICT 341 341 Q->H (IN REF. 2).
FT CONFLICT 528 528 V->A (IN REF. 2).
SQ SEQUENCE 630 AA; 73042 MW; 248A4B90319BC32D CRC64;
Query Match 26.6%; Score 1027.5; DB 5; Length 630;
Best Local Similarity 36.0%; Pred. No. 7, 7e-67;
Matches 215; Conservative 110; Mismatches 250; Indels 23; Gaps 9;
QY 6 QAKFELDKFNHEADLFYQSSLSAWNTNTEENVANNMN-AGDKMSAFKQSTLAOMY 65
DB 33 EAREFLELEQULRRRHEEFLGYNTNTVTAANROMAEVYARAEALKKRLAQIKSS 92
QY 66 PLQEQMLTVKQLQALQNGSSVSEDSKRLNTLTMTSTYTGKVNPNPQEC-L 124
DB 93 DYQSEADALRQAEHLISKAGSALNADIDIALQNAISSQNTYATAYVCSYTNRSQSL 152
QY 125 LLEPGLNEIANSIDYERLWAMESRSEVGKOLRLPYEEYVLKNGMAANYEYDGY 184
DB 153 TLEPHIOERLSHRDPAELMAYREWHKSGTBPONFAEYVRLTRKASQLNHSRYADY 212
QY 185 WRGDEYVNGVDYDSRGQLEDEYHTEFEIKLYEHLNAYRAKLNANAY-PSYISPIG 243
DB 213 WYQFYE-----DDFER-----QIDATFKQLLEPLRYQLHGYVFRRLQHGDPVMAEON 262
QY 244 LPAHLIGDMGRFTNLYSLTVPFGOKPNIDVDANVQAMQRIKFEAKETVAVGLP 303
DB 263 IPTSLIGNMWGQSMNELLDTFTPYPKPFVDYVAEMKQGYVYQKLFELGDPFGSLGR 322

QY 304 NMTGCFWNSMLTDCGNQKAVCHPTAMDGLK-GDFRLMCTKYTMDELTAHHMHQ 362
 DB 323 ALPSEFWMLSVLTTRDD-RQVVCASANDPFDQSDSVRKMKCTEVDSHFYVHHHLGHQ 381
 QY 363 YMAAFAOPFLIRNGANGFHEAVGEIMSLAATPKHLKSLGSLSPDFOEDNETEINFL 422
 DB 382 YLQYEQDPAYVRGAPNGFHEAVGDVIALSVMSAKHLKAGLIE-NORLDEKSRINOLF 440
 QY 423 KQALITVGTLPPTVYLEKRMWVFKGEIPKQDMKKMKMKREIVGVYVPPHDETCDP 482
 DB 441 KQALSKIVLPFGVADKYRAVFNELDESQMGNGFQMSSEFGVPPFRTKEDDP 500
 QY 483 ASLFFVNSDYSTRYRTTYRTRVFOFOALCOAKHEGP-----LHKDISNSTEAGOKL 536
 DB 501 PAKYHIDAVETLRFAHIFOFHAKLCRAAGYAPRNSRLTLNDNDIGSKRAAGSL 560
 QY 537 FNNLRIGKSEPTVLTALENVGAKNNVRLNYFEPLFTWLKDOKNSFVGNSTWSP 594
 DB 561 SQFLSGNSRHMKEVLEETGETEMDPALLLEYFPLVQWLKOE--NSRLGVPLGMP 616

RESULT 15
 Q9VLJ6 PRELIMINARY; PRT; 630 AA.

AC 09VLJ6: 01-MAY-2000 (TREMELREL. 13, Created)
 DT 01-MAY-2000 (TREMELREL. 13, last sequence update)
 DE 01-DEC-2001 (TREMELREL. 19, last annotation update)
 DE ACER protein (J028328F).
 GN ACER OR CG10593.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Phyllophaga; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer G., Champagne M., Pfeiffer B.L.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Bayandale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burks K.C., Busam D.A., Butler H., Cledon E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L., Domingues M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodde K., Gong F., Gorrell J.H., Gu Z., Guan P., Harris J.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hejlskov M., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
 RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laako P., Lei Y., Levitsky A.A., Li J., Li Z., Lian X., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhibo X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).

RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Partan D., Frise E., George R.,

RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,

RA Nunoo J., Paele J., Parag V., Park S., Phouanavong S., Wan K.,

RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,

RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AE003621; AAF52693.1;

DR EMBL: AY051750; AAK93174.1;

DR MEROPS: M02.002;

DR FLYBASE: FBgn001612; Acer.

DR InterPro: IPR001548; Peptidase_M2.

DR InterPro: IPR000130; Zn_MTPeptide.

DR Pfam: PF01401; Peptidase_M2; 1.

DR PRINTS: PR00791; PEPTIDASEA.

DR ProDom: PD004184; Peptidase_M2; 1.

DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN 1.

SQ SEQUENCE 630 AA; 73057 MM; 6D9355EB5773289 CRC64;

Query Match 26.5%; Score 1025.5; DB 5; Length 630;
 Best Local Similarity 36.0%; Pred. No. 1.1e-66;
 Matches 215; Conservative 109; Mismatches 251; Indels 23; Gaps 9;

QY 6 QAKTFLEKFNHEDLDFYQSSLAAMNTNTTEENYONMNNAGDKSAFIKEGSTLAOMY 65
 DB 33 EARRFELENQULRRRHEEFLSGYNTNTTEENYONMNNAGDKSAFIKEGSTLAOMY 92
 QY 66 PLEQIQLFVKIQLQALQNGSSVLSSEKSKRLTILNTSTIYSGKGNPNPQPC-L 124
 DB 93 DYQSESDIRQAEHLKSLGASALNMDVIALQNALISSQGTNAATVCSYTRSDCSL 152
 QY 125 LLEPLNEIMANSIDYNERLAMESWSSENGKOLRPLEEYVILKNMARNEYDGY 184
 DB 153 TLEPHIQRLSHSRDPAALAWYREWHDKSTPKNRPAEVRITLRAASQINGRSYADY 212
 QY 185 WRGDEVGVGVGYDVSRLGLEDVHPEEIKPLYEHLHAYVRALKNNAY-PSYISPIGC 243
 DB 213 WQGYFE-----DPPFER-----QIDATPKQLPYRQLHGVRRRLQOHGPDVMPAEGN 262
 QY 244 LPAHLIGDMGRFWNTIYSLTFPGQKPNIDVTDAVDAQRIKFAKEKFFVSGLP 303
 DB 263 IPISLIGMMWGSNNELDLFTPYEKEPFVDVKAEMEKGYVOKLELGDQFQSGMR 322
 QY 304 NMTGCFWNSMLTDCGNQKAVCHPTAMDGLK-GDFRLMCTKYTMDELTAHHMHQ 362
 DB 323 ALPSEFWMLSVLTTRDD-RHVVCHASANDPFDQSDSVRKMKCTEVDSHFYVHHHLGHQ 381
 QY 363 YMAAFAOPFLIRNGANGFHEAVGEIMSLAATPKHLKSLGSLSPDFOEDNETEINFL 422
 DB 382 YLQYEQDPAYVRGAPNGFHEAVGDVIALSVMSAKHLKAGLIE-NORLDEKSRINOLF 440
 QY 423 KQALITVGTLPPTVYLEKRMWVFKGEIPKQDMKKMKMKREIVGVYVPPHDETCDP 482
 DB 441 KQALSKIVLPFGVADKYRAVFNELDESQMGNGFQMSSEFGVPPFRTKEDDP 500
 QY 483 ASLFFVNSDYSTRYRTTYRTRVFOFOALCOAKHEGP-----LHKDISNSTEAGOKL 536
 DB 501 PAKYHIDAVETLRFAHIFOFHAKLCRAAGYAPRNSRLTLNDNDIGSKRAAGSL 560
 QY 537 FNNLRIGKSEPTVLTALENVGAKNNVRLNYFEPLFTWLKDOKNSFVGNSTWSP 594
 DB 561 SQFLSGNSRHMKEVLEETGETEMDPALLLEYFPLVQWLKOE--NSRLGVPLGMP 616

RESULT 16
 Q8SX2

ID 08SXX2 PRELIMINARY; PRT; 844 AA.
 AC 08SXX2: PRELIMINARY; PRT; 844 AA.
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE RH06639P.
 GN ANCE-3.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_taxid=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Cejner S.,
 RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AY075532; AAL68339.1;
 DR SEQUENCE 844 AA; 98243 MW; 46E1E123F5C8F71 CRC64;

Query Match 24.0%; Score 930; DB 5; Length 844;
 Best Local Similarity 33.2%; Pred. No. 1.7e-59;
 Matches 196; Conservative 119; Mismatches 248; Indels 24; Gaps 9;

QY 28 ASNNYNTNTEENVQNNAGDKMSAFLEQSTLAOMYPLDQIOMLYKLOLQALQNGS 87
 DB 234 AQNNFETNVNDFQTEALNAGQRYVEFGQITADQSRINKDILFRRLYRQLMLOSEVG 293
 QY 88 SYLSEKSKRLNITLNTASTYSTGVCVCPNDPQC-LILEFGLNEIMANSIDYNERLMA 146
 DB 294 NALPLDVLDRYNLEMLFLYNSAFCICAYQOPFCODLHYIQLKDIAKSDMDLQHT 353
 QY 147 MESWSEVQKOLPLTEEVVYLNEMARANHEDYGDWRGDYEVNGVDYDSGQILE 206
 DB 354 WVEYHKKAGRGKNDSEYGLIDMDEFAVYNNNTNGEWTYLAE-----SGNFRQ 403
 QY 207 DVNHTFEETKPLYLEHAAVYAKLMANAY-PSYISPGCLPAHLGDMMGRFWTMYLSLV 265
 DB 404 DMQIVNQIRPLIEGLHAYVRKRLDYQPRINRIAPISPHILGNMYGQSMNVLDILI 463
 QY 266 PEGQKPNIDVAMVQADQADQIRKPEKKEFVSGLMNTQGFRENSMLTDCGVQKAV 325
 DB 464 PYRGRKLIDVPRMVBQGTTPOLMQLAEEFETSNMSAVGPPEFYRNSIFRQPLD-RRVL 522
 QY 326 CHPTAMD-L-GKGFRLIMCTKVTMDFLTAHHEMHIQDYAAVAAOPLLRNGANEGPHE 384
 DB 523 CEPASAMDFCNRRHDFVYKICTDINOSLSVHHEMAHIOYFYQKHLKRIFFNGANPFRHQ 582
 QY 385 AVGEIMLSAATPKRLKSTIGLSPPFQDNTEINFLIKQALTJVTGLPFTYMLEKRW 444
 DB 583 AVGDAGIGSVSTPRKLTQTLGLQSRSLDE-SSYDINLYFTMAIDVAFALPALSIDNRWD 641
 QY 445 VKEGIEIPDQMKKMKWEKREIVGVPEYVPEDEYTCBPASLFHSNDYSFRYTRTLXQ 504
 DB 642 VESGANRRTMNCCHVMNRKESYSGKPPVYLSKEDDFGACATYHNPANIPKIFEFSTVQ 701
 QY 505 FQFQALCQAANK-----EGPLKCDISNSTEAGOKLFWMLRIGKSEPTWLLAENVGA 558
 DB 702 FOIYGLCRESGQYVGPDPKPKLHOCDIYRQPAAGNLIKFLTMSKGASQPOWELETLRE 761
 QY 559 KNNMVRPLNTEFPELFTWKLQDN--KNSPVGMSDMSPYADQSK 601
 DB 762 GRIDGTALREYFAFLLEMLRQENLRNRYEYGMND-CDYCKRSIE 805

RESULT 17
 Q9VJVI PRELIMINARY; PRT; 792 AA.

AC Q9VJVI:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE CG17988 protein.
 GN ANCE-3 OR CG17988.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_taxid=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY.
 RA MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Abell J.F., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Ballew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher S., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dunham K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fjosek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jellai M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Pelazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AEO03641; AAF5356.1;
 DR FlyBase: FBgn0032536; Ance-3.
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR000130; Zn_Mpeptidase.
 DR Pfam: PF01401; Peptidase_M2; 1.
 DR PRINTS: PR00791; PEPDIPASEA.
 DR ProDom: PD004184; Peptidase_M2; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 DR PROSITE: 792 AA; 92321 MW; 757EC47208902280 CRC64;

Query Match 21.7%; Score 841; DB 5; Length 792;
 Best Local Similarity 30.5%; Pred. No. 5e-53;
 Matches 196; Conservative 105; Mismatches 224; Indels 118; Gaps 14;

QY 30 MNVNTNTEENVQNNAGDKMSAFLEQGS--TLAQMT-----PLQSIQNLTVK 76
 DB 158 YNNNNNVEFV--GINNNRNFENFLSNDRFNLDGTERORYODRRYQGLTFLRTI 216
 QY 77 L-----QLQALQNGSSV 89

Db 217 LVSDQGSLECTNLNAQRIYEFQRTAEOGRKINKDIIPDRRLRLQMLQSEVPNA 276
 QY 90 LSEDKSKRLNTLNTSTSTYSGKVCNPDPQEC-LILEPGLNEIMANSIDYERLMAE 148
 Db 277 LPIDVLDREYNRLNEMFLVYSAEICAYQOPQCDLHYIPOLKIDIAKSDMDLQHTWV 336
 QY 149 SWRSVGVKQRLPLVEEYVVLKNEARANNHEDYGDYRGVYENGVDGYYSKGLIEDY 208
 Db 337 EYHRAAGKMDSEYQILDVQVEAYVNNNGEYLYAE-----SGNRQDM 386
 QY 209 EHTFEIKPLVLEHAAVYRAKIMNAV-PSYISPIGCLPAHLGDMGRTNLSLTPV 267
 Db 387 DIWQOIRPLVEGLHAAVYRRLRDYGGDRINRLAPISHLGIMVYQOSMNVLDILIPY 446
 QY 268 GQKPNIDVTAMDADQRIEKEAEKFEYVSLPMQGFENSLDTPGVYOKAYCH 327
 Db 447 PGRKLDIYTPRVVEGGYRQPLMFQLAEEFFTSIMSVAGPEFRKNSLFEQPLD-RVYLCE 505
 QY 328 PTAWPL-GKDPRLMCTKVTMDPLFAHHEMGHIQYDMAVYAQPELLRANGANGEPHEAV 386
 Db 506 PSAMPFCNRHDFRVAICTDINORSLSIVHMAHQYFLOGRHLPIFRGANGAPAFHQAV 565
 QY 387 GEIMSLSATPKRHLKISGLSPDQENETELNFKQALTYGTLEPTMLKERNMVF 446
 Db 566 GDAIGLSYSTRHLQTLGLQRLSIDE-SSYDINLYFWALDKAFLEFALSLOMWRDYVF 624
 QY 447 KEIIRKDDMMKMKWEMKREIYGVVPEPHDEYCDPASLFIHVSNDSEIRYETLYQFO 506
 Db 625 SGNAKRTYMNCHYMN-----LRFESTVLFQF 651
 QY 507 FOEALCQAANK-----EGPLHKDISNSTEAGOKLFNMLRLKSGSEPTLALENVGAKN 560
 Db 652 ITRGICRESGQYVPGDPRKPLHQCDYRQPAAGNLTKLMKSGASQPMQEVLEETLREGR 711
 QY 561 MNRVPLNTEPELFTWLQDN--KNSFGWSTDSVPYDOSIK 601
 Db 712 LGTALREYFAPLEMLELROENLRTEYVGWMD-GDYCKRSIE 753
 RESULT 18
 QY 99D836 PRELIMINARY; PRT; 249 AA.
 AC Q9D836;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE 20103051055Rik protein.
 GN 20103051055Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid-10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6J; TISSUE-SMALL INTESTINE;
 RX MEDLINE-21065660; Pubmed-11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arai K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamana A. I.,
 Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Giusti C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 Schirral L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Balderrelli R., Barsh G.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata K., Storch K.-F.,
 Suzuki H., Toyoko K., Wang K.H., Wetz T., Whitaker C., Wilming L.,
 Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK008530; BAB5723.1; -.
 DR MEROPS; M02.006; -.
 DR MGI; MGI:1917258; 20103051055Rik.
 DR InterPro; IPR001548; Peptidase_M2.
 DR Pfam; PF01401; Peptidase_M2; 1.
 SQ SEQUENCE 249 AA; 28379 MW; 19372B2878AAE921 CRC64;

Query Match 19.7%; Score 762; DB 11; Length 249;
 Best Local Similarity 78.0%; Pred. No. 5.9e-48;
 Matches 142; Conservative 20; Mismatches 20; Indels 0; Gaps 0;

QY 539 MLRLKSEPTLALENVGAKNNVRLPLNTEPELFTWLKQDNKSGVSTDSVPYADQ 598
 Db 1 MSLGNSSEPTKALENVGAKNNVRLPLNTEPELFTWLKQDNKSGVSTDSVPYADQ 60
 QY 599 SIKVRLSLSKSLGDKATKENDNEMKTLFRSSVAYAMRKYFSLTKVKNOMILGGEEDVRANKL 658
 Db 61 SIKVRLSLSKSLGDKATKENDNEMKTLFRSSVAYAMRKYFSLTKVKNOMILGGEEDVRANKL 120
 QY 659 PRISFVFYTPAKNVSDIIPTEVEKAIKMSRSHINDAFRLNDSLEFLGIQPTLPNQ 718
 Db 121 PAVSYFFVTBQNVSDVIRSEVEDAIRMSGRINOVFGUNDSEFLGIHPTLPYQ 180
 QY 719 PP 720
 Db 181 PP 182

RESULT 19
 QY 018581 PRELIMINARY; PRT; 907 AA.
 AC 018581;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE Hypothetical 101.1 kDa protein.
 GN C42D8.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_Taxid-6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE-99069613; Pubmed-9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology." The C. elegans Sequencing Consortium."
 RL Science 282:2012-2016(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX Hallsworth K.;
 RA "The sequence of C. elegans cosmid C42D8."
 RT Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX Waterston R.;
 RA "Direct Submission."
 RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; U56966; AAA98719.1; -.
 DR InterPro; IPR001548; Peptidase_M2.
 DR Pfam; PF01401; Peptidase_M2; 1.
 DR ProDom; PD004184; Peptidase_M2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 907 AA; 101086 MW; F9B8B5C0F9BC5MA3 CRC64;

Query Match 16.5%; Score 638.5; DB 5; Length 907;
 Best Local Similarity 27.0%; Pred. No. 4.2e-38;

Matches 167; Conservative 132; Mismatches 280. Total 579

[illegible]

RESULT 20

ID 09V520 PRELIMINARY; PRT; 661 AA.
 AC 09V520;
 DT 01-MAY-2000 (Tremblrel, 13, Created)
 DT 01-MAY-2000 (Tremblrel, 13, last sequence update)
 DT 01-JUN-2002 (Tremblrel, 21, last annotation update)
 DE CG8196 protein.
 GN ANCE-4 OR CG8196.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId:7227;
 [1]
 SEQUENCE FROM N.A.
 RP STRAIN-BERKELEY;
 RC MEDLINE=20196006; Pubmed=10731132;
 RX ADAMS M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananthanandis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazie R.G., Champe M., Peiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abriil J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck H.J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Cantor A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Duran-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gong F., Gough E., Harney D., Heiman T.J., Hernandez J.R., Houck J.,
RA Harris N.L., Harvey D., Horrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kratt C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., Meleod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris C., Mostrelli A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relnick K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sildenafil K., Stimpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stimpson M., Strong R., Sun E.,
RA Svitskas R., Teclor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weltsch G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye Y., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*."
RL EMBL: 287.2185-2195(2000).
DR EMBL: AE003835: AAF59002.1; -
DR FlyBase; FBgn003366; Ance-4;
DR InterPro; IPR001348; peptidase_M2.
DR Pfam; PF01401; peptidase_M2; 1.
DR PRINTS; PR00791; PEPDIP1ASEA.
DR PRODOM; PD004184; peptidase_M2; 1.
50 SEQUENCE 661 AA; 76218 MW; 2C0B768AC607306E CRC64;

Query Match	12.7%	Score 491;	DB 5;	Length 661;
Best Local Similarity	23.5%;	Pred. No. 1.7e-27;		
Matches 163; Conservative 133; Mismatches 309;				

missmatches 298; Indels 100; Gaps 25;

[illegible]

Tue May 27 09:36:36 2003

us-09-978-385-2_copy_19_738.rspt

Page 14

```

OY      453  DQPMKFMWEMKEKREIVGVEPVPHEDTTCYCDPASLPHVSNDSIF--IRYTRTRILYQFOFOEAL 511
Db      462  DYNKHKHYWRLMEQHOAGIEAPSDRSFGALIDFPYKRYVYNDSEFOFOKITEVLGYOFYREE 521
OY      512  COAAKHGEPRLHCODISNPFEGACOFENMLRLGKSEPTLALENVGA----KNMNVRL 566
Db      522  CKSYNKGPLHNCDFYGLSLAVGNDNDKAMMSLGSTFPW----KQVYVKILPNTGLSIAL 577
OY      567  LNTFEPFLWLKQDN--NSFQWGSIDSWEPYADOSIKYRISLSKSLGDKAEYNDNDMYL 624
Db      578  LEYQOPFLDMLNTNYNDANSKIGWIA-----SKMAYVT-----GGDDRLYF 618
OY      625  FRSSVAYAMROYELFKYKNOMILLEGEDVEVANIK 658
Db      619  AHVYDTLSLNSVMYVIRN--VYDSFENIKYILO 649

```

Search completed: May 26, 2003, 17:55:42
Job time : 73 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 17:58:24 ; Search time 4974 Seconds

(without alignments)
4212.710 Million cell updates/sec

Title: US-09-978-385-2_COPY_19_738

Sequence: 3869
1 STRLEQAKFLLDKFNFHEAD.....DNSLEFLGIQPLGPPNPQP 720

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 205640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O/cgn2.1/USPTO.spool/US09978385/runat_21052003_154419_19834/app_query.fasta_1.903
-DB=GenB1 -GFWT=fastap -SUFFIX=eye -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=20 -MODE=LOCAL
-OUTPRT=Pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09978385.ecgn_1.1.3284.gunat_21052003_154419_19834 -NCPU=6 -ICPU=3
-NO_MAP -LARGOQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELext=7

Database :

GenB1.*
1: gb_da.*
2: gb_hlg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_ot.*
22: em_ov.*
23: em_pat.*
24: em_pl.*
25: em_pi.*
26: em_ro.*
27: em_sts.*
28: em_un.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3869	100.0	2415	6 AR135178	AR135178 Sequence
2	3869	100.0	2415	6 AX418984	AX418984 Sequence
3	3869	100.0	2415	6 E43987	E43987 ACE-analogo
4	3869	100.0	2418	6 E39033	E39033 MPOR15 pol
5	3869	100.0	2599	6 E43988	E43988 ACE-analogo
6	3869	100.0	3325	6 AF291820	AF291820 Homo sapi
7	3869	100.0	3334	6 AX047758	AX047758 Sequence
8	3869	100.0	3341	6 HSM800880	AL110224 Homo sapi
9	3869	100.0	3396	6 AR135177	AR135177 Sequence
10	3869	100.0	3396	6 AX418982	AX418982 Sequence
11	3869	100.0	3405	6 AX421254	AX421254 Homo sapi
12	3869	100.0	3405	6 AB046569	AB046569 Homo sapi
13	3866	99.9	2599	6 AB046569	AB046569 Homo sapi
14	3720	96.1	3732	6 E39034	E39034 MPOR15 pol
15	3366.5	87.0	2262	6 E39034	E39034 MPOR15 pol
16	3251	84.0	2638	6 AX047762	AX047762 Sequence
17	3251	84.0	2739	10 BC026801	BC026801 Mus muscu
18	3251	84.0	2760	10 AB053181	AB053181 Mus muscu
19	3233	83.6	2638	6 AX047765	AX047765 Sequence
20	3175	82.1	2415	6 AX047760	AX047760 Sequence
21	2645	68.4	2415	6 AX047764	AX047764 Sequence
22	1844	47.7	1993	10 AB053182	AB053182 Mus muscu
23	1345.5	34.8	4879	9 BC036375	BC036375 Homo sapi
24	1337	34.6	2473	9 HUMRACER	M2657 Human testl
25	1337	34.6	2477	9 HUMRACER	X16295 Human testl
26	1337	34.6	2478	6 A31567	A31567 H.sapiens t
27	1337	34.6	4020	6 AR037213	AR037213 Sequence
28	1337	34.6	4020	6 AR137383	AR137383 Sequence
29	1337	34.6	4020	9 HUMRACER	J04144 Human angio
30	1337	34.6	4022	6 AX429555	AX429555 Sequence
31	1337	34.6	4024	6 A00914	A00914 H.sapiens g
32	1334	34.5	2418	10 MUSACER	M55333 Mouse testl
33	1334	34.5	3681	10 MUSACER	BC034367 Mus muscu
34	1334	34.5	3813	10 MUSACER	J04946 Mouse angio
35	1334	34.5	4694	6 AX147503	AX147503 Sequence
36	1334	34.5	4694	10 MUSACER	AX147503 Sequence
37	1334	34.5	2551	10 MUSACER	J04947 Mouse angio
38	1315	33.9	4050	5 CHKACEI	J04947 Mouse angio
39	1312	33.9	4014	10 AF201331	M29981 Human abert
40	1310	33.9	4014	10 AF201332	L40175 Gallus gall
41	1310	33.9	4014	10 AF201332	AF201332 Rattus no
42	1310	33.9	4014	10 RND03708	AE201332 Rattus no
43	1303	33.7	2490	9 RND03734	U03708 Rattus norv
44	1303	33.7	2490	9 HUMRACER	U03734 Rattus norv
45	1283	33.2	2409	4 RABRACA	M2658 Human testl

RESULT 1

ALIGNMENTS

AR135178 2415 bp DNA linear PAT 16-MAY-2001
 LOCUS AR135178 3 from patent US 6194556.
 DEFINITION AR135178
 ACCESSION AR135178.1 GI:14124083
 VERSION AR135178.1
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2415)
 AUTHORS Acton, S. Laurene, and Robison, K. Earl.
 TITLE Angiotensin converting enzyme homolog and therapeutic and
 diagnostic uses therefor
 JOURNAL Patent: US 6194556-A 3 27-FEB-2001;
 FEATURES
 source location/Qualifiers
 1. 2415
 /organism="unknown"
 BASE COUNT 743 a 483 c 555 g 634 t
 ORIGIN

Alignment Scores:
 pred. No.: 0 Length: 2415
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AR135178 (1-2415)

OY 1 SerThrIleGluGluGluAlaIleThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 55 TCACCATGAGACAGACAGCCAGCAATTTTGGACAGTTAACCCAGAGCCGAGAC 114
 OY 21 LeuPheTyrGlnSerSerLeuAlaSerThrPasnTyrAsnThrAsnIleThrGluGluAsn 40
 DB 115 CTCTCTATCAAAAGTCACTTCTCTTCTGGAATTTTAAACCAATATTCTGAAGAGAT 174
 OY 41 ValGlnAsnMetAsnAsnAlaGluAspLysTrpSerAlaPheLeuGluGlnSerThr 60
 DB 175 GTCCAAAACATGATATATGCTGGGACCAAAATGCTGCTTTTAAAGAACAGTCCACA 234
 OY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 235 CTGGCCCAATGTATCCACTACACAAATTCAGATCCACAGTCAAGCTTCAGCTGAG 294
 OY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 295 GCTCTTCAGCAAAATGGCTCTTCAGTGTCTGAGAAACAGAACAGCAACGTTGAACACA 354
 OY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 DB 355 ATTCTAATATCAATAGACACCACTACAGTACGAAAGTTTGTAAACCAATATATCCA 414
 OY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 415 CAAACAATGCTTATTAATCTTGAACCAAGTTTAAATGAATATATGCAACAGTTTAAGCTAC 474
 OY 141 AsnGluArgLeuThrPalaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
 DB 475 AATGAGAGGCTCTGGGCTGGGAAAGCTGGAGATCTGAGTCCGCAAGCAGCTGAGGCCA 534
 OY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 535 TTATATGAAGAGTATGCTGCTTGAATAAATGAGTGGCAAGAGCAAAATCTTTATGAGAC 594
 OY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 DB 595 TATGGGAGTATTTGAGAGGAGACTATCAATTAATGGGTAAGATGGCTATGACTACAGC 654
 OY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 DB 655 CCGGCGCAGTGTATTTGAAGATGTGAGACATACCTTGTGAAGATTAACCATATATATGAA 714

OY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 715 CATCTTCATATCCATATGAGAGGCAAAAGTTGATGATGCTATCTCTATATATGCTCA 774
 OY 241 IleGlyCysLeuProAlaHisLeuGluGlyAspMetTrpGlyArgPheTrpThrAsnLeu 260
 DB 775 ATTGAGTCCCTCCCTGCTCATTTGCTGGTGAATGTGGGTGATTTTGGCAAAATTTG 834
 OY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 835 TACTCTTGAACAGTCCCTTGGACGAAACCAAAATGATGATGATGATGATGATG 894
 OY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 895 GACCAAGCCTGGATGACAGAGAAATTCAGAGAGCCGAGAGAGTTCTTGTATCTGT 954
 OY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 955 GGTCTTCTATATGACTCAAGATTTCTGGAAATTTCCATGTTAACGACCCAGCAAAAT 1014
 OY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 DB 1015 GTTCAGAAAGCAGTCTGCCATCCACAGCTGGACCTGGGAGGCGCACTTCAGAGAC 1074
 OY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHis 360
 DB 1075 CTATGTCCACAAAGCTACATGATGACACTCTCTGACAGCTCATCATGAGATGGGCAT 1134
 OY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1135 ATCCAGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 1194
 OY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaIleThrProLysHisLeu 400
 DB 1195 GATTCATCAAGAGCTGTGGGAAATTCATGCACTTCTGACGCCACACCTTAACATTA 1254
 OY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 DB 1255 AATTCATGCTGCTTCTGACCCGATTTTCAAGAGACATTAAGAAACAGAAATTAACCTC 1314
 OY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1315 CTGCTCAAAACAGACATCAGATGTGGAGTGTGCACTTACTTACATTTTGAAGAG 1374
 OY 441 TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrp 460
 DB 1375 TGGAGGTGATGTCTTTAAAGGAAATTCACCAAGAGTGAATGAAGAGTGGTG 1434
 OY 461 GluMetLysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCys 480
 DB 1435 GAGATGAAGCGAGATATGATGGGTGGTGGAAACCTGTCACATGATGAAGACATCTGT 1494
 OY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrArg 500
 DB 1495 GACCCGCACTCTGTTCCATGTTTAAATGATTAATCTATCTATGATATACCAAGG 1554
 OY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 DB 1555 ACCCTTACCAATTCAGTTTCAAGAAAGCACTTTCAGAGCAGCTAAACATGAAGCCCT 1614
 OY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGluLysLeuPheAsnMetLeu 540
 DB 1615 CTGCACAAATGTGACATCTCAAACTCTACAGAAAGCTGGACAGAAAGCTGTCAATATGCTG 1674
 OY 541 ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsn 560
 DB 1675 AGGCTTGAATAATGAAGAACCTGAGACCTGCAATGTGAATAATTTGTAGAGCAAGAAC 1734
 OY 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrPheLysAspGln 580
 DB 1735 ATGATGTAAAGCACATGCTCACTTGTAGAGCCCTTATTTACCTGGCTGAAGACACAG 1794

QY 581 AsnlysnserPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIle 600
 Db 1795 AACAGAAATCTTTTGGATGGATGAGTACCGACTGGATCCATATGACGACCAAGCATC 1854
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrAsnAspAsn 620
 Db 1855 AAGTAGAGTAAAGCTTAAATCAGCTCTTGGAGATTAAGCATATGAAATGGAACGACAAAT 1914
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuVal 640
 Db 1915 GAATGTACCTGTCGATCATCTGTGCATATGATATGAGGCACTATTTTAAAGTA 1974
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGlyAspValArgValAlaAsnLeuLysProArg 660
 Db 1975 AAAATCATGATATCTTTTGGGAGGAGATGTCGAGCGCTAAATTTGAAACCAAGA 2034
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 Db 2035 ATCTCTTAAATTTCTTGTCTACATGACCTTAAATATGTCGTATCATCTTCTAGAACT 2094
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 Db 2095 GAAGTTGAAAGGCGCATAGCATGTCGCGAGCGGTATCATGATGCTTCCGTCGAAT 2154
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2155 GACAAACACCTAGAGTTCTGTGGGATACAGCAACACTTGGACCTCTTAACGACGCCCT 2214
 RESULT 2
 AX418984 2415 bp DNA linear PAT 18-JUN-2002
 LOCUS Sequence 3 from Patent WO0212471.
 DEFINITION AX418984
 ACCESSION AX418984
 VERSION AX418984.1 GI:21523783
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Acton, S., Robison, K.E. and Hsieh, F.Y.
 TITLE Angiotensin converting enzyme homolog and uses therefor
 JOURNAL Patent: WO 0212471-A 3-14-FEB-2002;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source 1..2415
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 743 a 483 c 555 g 634 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2415
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AX418984 (1-2415)
 QY 1 SerThrIleGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 Db 55 TCCACCATTTGAGGACAGCCAGCATTTTGGACAAAGTTTAAACCGAAGCCGAGAGAC 114
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrAsnTyrAsnThrAsnIleThrGluLysAsn 40
 Db 115 CTGTTCTATCAAAAGTCTCTTCTTGGAAATTAACACCAATATTTACTGAAGAGAAAT 174
 QY 41 ValGlnAsnMetAsnAlaGlyAspLysTyrSerAlaPheLeuLysGluGlnSerThr 60
 Db 175 GTCCAAAACATGATATATGCTGGGACAAATGCTCTCTTTTAAAGAACAGTCCACA 234

QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 Db 235 CTGGCCAAATGATATCATCTACAGAAATTCGAATCTTCACAGTCAACCTTCAGCTGAC 294
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 Db 295 GCTCTTACCAAAATGAGTCTTCAGTGTCTGAGAAAGCAAGAGCAAAACGTTGACACA 354
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 Db 355 ATCTTAATACATATGAGCCACCTTACAGTACTGGAAAAAGTTTGAACCCAGATAATCCA 414
 QY 121 GlnGluCysLeuLeuLeuGluLeuProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyr 140
 Db 415 CAAGAAATGCTATTTACTTGAACACAGCTTGAATGAATTAATGCCAAACAGTTTATGACTAC 474
 QY 141 AsnGluArgLeuThrAlaTyrGlnSerTyrPheArgSerGluValGlyLysGlnLeuArgPro 160
 Db 475 AATGAGAGGCTCTGGGCTTGGAAAGCTGGAGATCTGAGGTGGGCAAGCAGTGAAGCCA 534
 QY 161 LeuTyrGluGluTyrValValIleLysAsnGlnMetAlaArgAlaAsnHisTyrGluAsp 180
 Db 535 TTATATGAAGAGTATGCTGCTTGAAAAATGAGATGGCAAGCAACATTAATGAGAGAC 594
 QY 181 TyrGluAspTyrTyrPheArgLysAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 Db 595 TATGGGATTTATGAGAGGAGACTATGAATGAATGGGTAGATGGCTATGACTGACTACAC 654
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGlnGluIleLysProLeuTyrGlu 220
 Db 655 CGCGGCGCAGTTGATGAAGATGAGCAATACCTTTGAGAGATTAACCATTAATATGA 714
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 Db 715 CATCTTATGCTATGATGAGGAGGCAAGTTGATTAATGCTTATCTTCTATATGACTCA 774
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheThrAsnLeu 260
 Db 775 ATTGAGATGCTCCCTGCTCATTTGCTGGAGATGCGGATGGGATTTGGCAAAATG 834
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 Db 835 TACTCTTGGACAGTCTCTTGGACAGAAACCAACATGATGATACATGATGCAATGCTG 894
 QY 281 AspGlnAlaTyrAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 Db 895 GACCAAGGCGGGATGCACAGAAATATTCAGAGAGCGGAGAACTTTGTATCTGCTT 954
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrGluAsnSerMetLeuThrAspProGlyAsn 320
 Db 955 GGTCTTCTTAATATGATCTCAAGATTTCTGGAAATTCATGCTAACCGACCCAGGAAT 1014
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIle 340
 Db 1015 GTTCAGAAAGAGTGTGCGATCCACACACTTGGGAGCTGGGGAAGGCGACTTCAGGATC 1074
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHis 360
 Db 1075 CTTATGTGCACAAAGTACAAATGACATGCGACTTCCGACAGCTCATGATGATGGGCAT 1134
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuMetArgAsnGlyAlaAsnGlu 380
 Db 1135 ATCCAGTATGATATGATGATATGCTGACACACCTTTTCTGTAACAAATGAGCAATATGA 1194
 QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisIleu 400
 Db 1195 GATTCATCAATGAAGCTGTGGGAAATCATGCTTCTGAGGACACCTTAAGCATTTTA 1254
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 Db 1255 AAATCATTTGCTCTTCTGTCACCGGATTTTCAAGAAACATATAACAGAAATTAACCTTC 1314
 QY 421 LeuLeuLysGlnAlaLeuThrIleValIleGlyThrLeuProPheThrTyrMetLeuGluLys 440

Alignment Scores:

Pred. No.:	Length:	2418
Score:	3869.00	720
Percent Similarity:	100.00%	0
Best Local Similarity:	100.00%	0
Query Match:	100.00%	0
DB:	6	0

US-09-978-385-2_COPY_19_738 (1-720) x E39033 (1-2418)

1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 55 TCACCATTTGAGGAGACAGGACATTTTGGACAGATTTAACAGCAGAGCGAGAGAC 114
 21 LeuPheTYGlnSerSerLeuAlaSerTrpAsnTrpAsnTrpAsnTrpAsnTrpAsn 40
 115 CTGCTCTAATCAAGCTTCACTGCTTCTTGAATTAACACCAATATTTACTGAGAGAAAT 174
 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
 175 GTCCAAACATGATTAATGCTGGGCAAAATGCTGCTTTTAAAGAAACAGTCCACA 234
 61 LeuAlaGlnMetLysTrpLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 235 CTGGCCCAATGATATCACTACACAGAAATTCAGAAATTCACAGTCAAGCTTCACTGAG 294
 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 295 GCCTCTCAGCAAAATGGCTTCAGGCTTCAGAGACAGAGCAAGCAAGCGTTGACACA 354
 101 IleLeuAsnThrMetSerThrIleTrpSerThrGlyLysValCysAsnProAspAsnPro 120
 355 ATCTCTAAATACATGAGCAACATCTACAGTACTGAGAAAGTTTGAACCCAGATTAATCCA 414
 121 GlnGluCysLeuLeuLeuGlnLuproGlyLeuAsnGluIleMetAlaAsnSerLeuAspTrp 140
 415 CAAGAATGCTTATTACTTGGACCAAGTTGAATGAATATGAGCAAGCAAGTTTGAAGTAC 474
 141 AsnGluArgLeuTrpAlaTrpGlnSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
 475 AATGAGAGGCTTGGGCTTGGGAAAGCTGAGATCTGAGGTGGCAAGAGCTGAGGCCA 534
 161 LeuTrpGluGluTrpValValLeuLysAsnGluMetAlaArgAlaAsnHisTrpGluAsp 180
 535 TTATATGAGAGATGATGGCTTGAAGAAATGAGATGAGCAAGCAAGCAATATATAGGAC 594
 181 TyrGlnAspTrpTrpArgGlyAspTrpGluValAsnGlyValAspGlyTyrAspTrpSer 200
 595 TATGGGATTAATGGAGAGAGAGACTATGAAGTAAATGGGTATGATGCTATGACTACAGC 654
 201 ArgGlyLysLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTrpGlu 220
 655 CGGCGCAGTTGATGATGAGATGTGGAACATACCTTGAAGAGATTAACCATTAATATGAA 714
 221 HisLeuHisAlaTrpValArgValAlaLysLeuMetAsnAlaTyrProSerTrpIleSerPro 240
 715 CATCTCAATGCTTAATGAGAGAGAGCAAGTTATGAATGCCATCTCCATATATGATGCA 774
 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpPheAsnLeu 260
 775 ATGTGATGCCCTCCCTGCTCATTTGCTTGGATATGAGGTATGAGATTTGGCAAAATCG 834
 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValAlaHisAlaMetVal 280
 835 TACTCTTGAACAGTCTCCCTTGGACAGAAACCAACCTAGATCTTACTGATGCAATGCG 894
 281 AsnGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 895 GACCAAGGCTTGGATGACAGAGAAATATCAAGAGGCGGAGAAATCTTGTATCTGTT 954
 301 GlyLeuProAsnMetTrpGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
 955 GGTCTTCTAATATGACTCAAGATCTCTGGGAAATTCATGTCTAAGGAGCCAGGAAAT 1014

321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 1015 GTTCAGAAAGAGCTCTGCAATCCACAGACTTGGAGCTGGGAGAGGCGACTTCAGAGATC 1074
 341 LeuMetCysTrpLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHis 360
 1075 CTTATGTCCAAAGAGTACCAATGAGAGACTTCTGACACCTCATCATGATGATGGGCAAT 1134
 361 IleGlnTrpAspMetAlaTrpValAlaGlnProPheLeuLeuTrpAsnGluAlaAsnGlu 380
 1135 ATCCAGTATGATGAGCATATGCTGCAACACTTTTGTCTAAGAAATGAGAGTATGAA 1194
 381 GlyPheHisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 1195 GGATTCATGAAAGCTGTTGGGAAATCATGCTTCTGACAGCCACACCTTAAGCAATTTA 1254
 401 LysSerIleGlyLeuLeuSerProAspPheGlnLysAsnGluTrpGluIleAsnPro 420
 1255 AATCCATTTGGCTTCTGCTCACCCGATTTCAAGAAACATGAAACAGAAATTAATCTTC 1314
 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTrpMetLeuGluLys 440
 1315 CTGCTCAAAACAGACATCAGATTTGTTGGACTGTCATTTACTTACATGTTAAGAGAG 1374
 441 TrpArgTrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysTrpTrp 460
 1375 TGGAGGTGATGCTGTTTAAAGGGAAATCCCAAGACCCAGTGAATGAAGAAATGGTGG 1434
 461 GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluTrpCys 480
 1435 GAGATACAGCAGAGATGATTTGGGTGGAGACTGAGCCATGATGAATCACTACTGT 1494
 481 AspProAlaSerLeuPheHisValSerAsnAspTrpSerPheIleArgTrpTrpArg 500
 1495 GACCCGCACTCTGCTTCATGTTCTATGATGATTAATCTCATTCATTCATGATTAACAG 1554
 501 ThrLeuTrpGlnPheGlnPheGlnLysAlaLeuCysGlnAlaAlaLysHisGluPro 520
 1555 ACCCTTACCAATTCAGATTTCAAGAGACATTTTCAAGCAGCAATGAACATGAAGGCCCT 1614
 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyLysLeuPheAsnMetLeu 540
 1615 CTCACAAATGTACATCTCAAACTCTACAGAGCTGAGCAAGAACTTTCAATATGCTG 1674
 541 ArgLeuGlyLysSerGluProTrpPheIleAlaLeuGluAsnValAlaGlyAlaLysAsn 560
 1675 AGGCTTGAAGAAATCAAGACCTGGACCTTACGATTTGGAAGAAATGTTGTAGAGCAAGAAC 1734
 561 MetAsnValArgProLeuLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAspGln 580
 1735 ATGATGTAAAGCCACCTGCTCAACTTGTGAAGCCCTTATTTACCTGCTGTAAGACACAG 1794
 581 AsnLysAsnSerPheValGlyTrpSerTrpAspTrpSerProTyrAlaAspGlnSerIle 600
 1795 AACAGAAATCTTTTGGGATGAGTACGAGTGGAGTCCATATGACAGCAAGCAATC 1854
 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn 620
 1855 AAATGAGATTAAGCCCTAAATACAGCTTGGAGATGAAGCAATGAATGAAGCAACAT 1914
 621 GluMetLysLeuPheArgSerSerValAlaTyrAlaMetArgGlnTrpPheLeuLysVal 640
 1915 GAAATGTACCTGTTCCATCATCATGTTGATGATGATGATGATGATGATGATGATGAT 1974
 641 LysAsnGlnMetIleLeuPheGlyGluAlaAspValArgValAlaAlaAsnLeuLysProArg 660
 1975 AAAATCAAGATGATCTTTTGGGAGAGAGATGTCGAGTGGCTAATTTGAACCAAGA 2034
 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 2035 ATTCCTTATATTTCTTGTGACAGCACTAAAGAAATGTGTGATATCATTTCTTGAAGCT 2094

QY 681 GLUVALGLULYSALALILEARGMETSERARGSERAGILEASNSPALAPHEARGLEUASN 700
 DB 2095 GAGGTTCGAAAGAGCCCTCAGATGTCGCCGAGCCCTATCAATGATGCTTCCGTCGAT 2154
 QY 701 ASPASNSERLIENGLUPHEUGLIYILEGINPROTHIRLEUGLIYPROPOASNGINPROPRO 720
 DB 2155 GACACAGCCTAGAGTTCTGTGGGATACAGCCACCTTGACCTCTTACACAGCCCT 2214

RESULT 5
 E43988
 LOCUS E43988 2599 bp DNA linear PAT 31-JAN-2002
 DEFINITION ACE-analogous gene.
 ACCESSION E43988.1 GI:18629191
 VERSION JP 2001046072-A/2.
 KEYWORDS unclassified.
 SOURCE unclassified.
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 2599)
 AUTHORS Sugano, S. and Komatsu, T.
 TITLE ACE-analogous gene.
 JOURNAL Patient: JP 2001046072-A 2 20-FEB-2001;
 OTSUKA PHARMACEUT CO LTD
 COMMENT OS Unknown
 PN JP 2001046072-A/2
 PD 20-FEB-2001
 PE 06-AUG-1999 JP 1999223892

PI SUMIO SUGANO, TAKAMI KOMATSU
 PC C12N15/09, A61K31/00, A61K31/7088, A61K38/00, A61K38/55,
 A61K39/395, PC A61K39/395,
 PC A61K39/395, A61K48/00, A61P9/12, C07K14/47, C07K16/08, C12N1/15, PC
 C12N1/19,
 PC C12N1/21, C12N5/10, C12Q1/68, G01N33/53, C12N15/00, A61K37/02, PC
 A61K37/64,
 PC C12N5/00
 CC
 FH
 FT

FEATURES
 source 1..2599 Location/Qualifiers
 /db_xref="taxon:32644"
 BASE COUNT 807 a 503 c 591 g 698 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 2599
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x E43988 (1-2599)

QY 1 SerThrIleGluGluGlnAlaIleLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 109 TCCACCATTCAGAGAACGCCAAGACATTTTGGACAAGTTTAACACGAGAACGCCAAGAC 168
 QY 21 LeuPheTyGlnSerSerLeuAlaSerThrPheLeuAsnHisGluGlnAsn 40
 DB 169 CCGTCTTATCAAGTTCACCTTCCTCTGGAATTTAAACACCAATATTCAGAACAGCAT 228
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
 DB 229 GTCCAAACATGAAATGCTGGGACAAATGGTCTGCTTTTAAAGGAACATCCACA 288
 QY 61 LeuAlaGlnMetTyPheLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 289 CTTCGCCAATATGATCCATCAACAAATTCAGAAATTCACAGCTTCAGCTGCAG 348
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThr 100

DB 349 GCTCTTCAGCAAAATGGGCTTCATGCTTCAGAGAACAGCAAGCAAGCTTACACACA 408
 QY 101 IleLeuAsnThrMetSerThrIleTySerThrGlyLysValCysAsnProAspAsnPro 120
 DB 409 ATTCATTAATACATAGAGCAACCATCTACAGTACTGGAAGAAATTTGTAACCCAGATATCCA 468
 QY 121 GlnGluCysLeuLeuLeuGluLeuProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTy 140
 DB 469 CAGAAATGCTTATCTTCAACACAGCTTGAATGAATGAATGAATGAATGAATGAATGAAT 528
 QY 141 AsnGluArgLeuThrAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuAspPro 160
 DB 529 AATGAGAGCTTGGGCTTGGGAAAGCTGGAGATCTGGGCTGGGCAAGCTGAGGCCA 588
 QY 161 LeuTyGluGluTyValValLeuLysAsnGluMetAlaArgAlaAsnHisTyGluAsp 180
 DB 589 TTATATGAAGAGTATGTGCTGTGAAAATGAGATGGCAAGCAATTCATATGAGAGC 648
 QY 181 TyGlyAspTyTrpTrpArgGlyAspTyGluValAsnGlyValAspGlyTyAspTySer 200
 DB 649 TATGGGATTTATGGAGAGAGACTATGAATGAATGGGTAGATGGCTATGACTACAG 708
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluIleLysProLeuTyGlu 220
 DB 709 CGCGCCAGTTCATGTAAGAGATGGAAACATACCTTTGAAGATTAACCATTAATATGAA 768
 QY 221 HisLeuHisAlaTyValValArgAlaLysLeuMetAsnAlaTyProSerTyIleSerPro 240
 DB 769 CATCTTCATGCTTATGAGAGGCAAAAGTTGATGATCCCTATCATATCATCA 828
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGluAspMetTrpGlyArgThrPheAsnLeu 260
 DB 829 ATFGATGCTCCCTGCTCATTTGCTTGGATATGGGGTAGATTTGGACAATTCG 888
 QY 261 TySerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 889 TACCTTTGACAGTTCCTTGGACAGAAACCAACATGATTTCTATCATCATGCTG 948
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPheValSerVal 300
 DB 949 GACCAAGCCTGGGATGACAGAAATATTCAGAGAGCCGAGAAATTTCTTATCTGCT 1008
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProLysAsn 320
 DB 1009 GGTCTTCCTAATATGACTCAAGATTCCTGGAAAATTCATCTTAACGACCCAGGAAT 1068
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 DB 1069 GTTCAGAAACAGTCTGCCATCCACAGCTGGAGCTGGGAGGGGAGCTTCAGATC 1128
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetCysHis 360
 DB 1129 CTTATTCGACAAAGGTGCAATGACGACCTTCGACAGCTCATCAATGAGATGGGAT 1188
 QY 361 IleGlnTyAspMetAlaTyValAlaGlnProPheLeuLeuArgAsnGlyValAsnGlu 380
 DB 1189 ATCCAGTATGATATGACATATGCTGCACAACTTTTTCGTAAGAAATGAGCATATGAA 1248
 QY 381 GlyPheHisGluAlaValGluIleLysSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1249 GATTCATCAAGAGCTTGGGAAATCATGCTTCTTGAGCCACCTTAACCATTTA 1308
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAsnGluIleThrGluIleAsnPhe 420
 DB 1309 AATATCATATGCTTCTGACCCGATTTTCAAGAGCAATGAAACGAATTAACCTTC 1368
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyLysLeuProPheThrTyMetLeuGlnLys 440
 DB 1369 CTTCCTCAACAGCACTACAGATTTGGACTCTGCCATTTACTTACATGTTAGAGAG 1428
 QY 441 TrpArgTrpMetValPheLysGlyLysLeuProLysAspGlnTrpMetLysLysTrpTrp 460

Db	1429	TGGAGGTGATGGCTCTTTAAAGGGGAAATTC	CCAAAGACAGTGATGAAAAAGTGGTG	1488	
OY	461	GIuMetLysrGluIleValGlyValGluProVal	ProHisAspGluThrTyrCys	480	
Db	1489	GAGATGAAAGGAGAGATAGTGGGGTGGCGAAC	CTGTGCCCATATGAAACATACCTGT	1548	
OY	481	AspProLAsSerLeuPheHisValSerAsnAsp	TyrSerPheIleArgTyrTyrThrArg	500	
Db	1549	GACCCCGCATCTCTGTTCATCTTTCTTAATG	ATTACTCATTCACTTGATATTCACAAAG	1608	
OY	501	ThrLeuTyrGlnPheGlnPheGlnGluAlaLeu	CysGlnAlaAlaLysHisGluGlyPro	520	
Db	1609	ACCCATTACCAATTCACAGTTTCAAGAGCACT	TGTTCACACACTAAACCTGAAAGGCCCT	1668	
OY	521	LeuHisTyrCysAspIleSerAsnSerThrGlu	AlaIleGlnLysLeuPheAsnMetLeu	540	
Db	1669	CTGCACAAATGTGCATCTCAAACTCTACAG	AAACCTGGACAGAAACCTTCATATGCTG	1728	
OY	541	ArgLeuGlyLysSerGluProTyrThrLeuAla	LeuGluAsnValValGlyAlaLysAsn	560	
Db	1729	AGGCTTGGAAATACAGAACCTTGACCTTGAC	TTCAGAAATGTGTAGGACGAAACAAAC	1788	
OY	561	MetAsnValArgProLeuLeuAsnTyrPheGlu	ProLeuPheThrTyrPheLysAspGln	580	
Db	1789	ATGATATGAAAGCCACATGCTCACTACTCTG	AGCCCTTATTACCTGGCTGAAAGACAG	1848	
OY	581	AsnLysAsnSerPheValGlyTyrPserThrAsp	TyrPserProTyrAlaAspGlnSerIle	600	
Db	1849	AACAAGATTCCTTTTGTGGGATGGAGTACCA	CGACATGCAATGCACAAACAGCATC	1908	
OY	601	LysValAlaGllIleSerLeuLysSerAlaLeu	GlnLysAspLysAlaTyrGluTyrPAsnAsp	620	
Db	1909	AAAGTGAGAGATTAAGCTTAAATTCAGCTTC	TGGAGATTAAGCATATGATGACAGACAT	1968	
OY	621	GluMetTyrLeuPheArgSerSerValAlaTyr	AlaMetArgGlnTyrPheLeuLysVal	640	
Db	1969	GAATAGTACCTCTTCCGATTCATCTGTTCAT	ATAGCTATGAGGACGACTTTTAAAGATA	2028	
OY	641	LysAsnGlnMetIleLeuPheGlyGluGluAsp	ValArgValAlaAsnLeuLysProArg	660	
Db	2029	AAAAATCAGATATTCCTTTTGGGAGGAGGAT	GTGGCGAGTGGCGAGTGGCTTAATTTGAA	2088	
OY	661	IleSerPheAsnPhePheValThrAlaProLys	AsnValSerAspIleIleProArgThr	680	
Db	2089	ATCTCTTTAATTTCTTTCTTCTCAGCTAA	AAATGATGCTCGATATTCCTTGAGACT	2148	
OY	681	GluValGluLysAlaIleArgMetSerArgIle	AsnAspAlaPheArgLeuAsn	700	
Db	2149	GAAGTGAAAGGCCATCAGATGTCGCCGAC	CCGTATCAATGATGCTTCCGTGATAT	2208	
OY	701	AspAsnSerLeuGlnPheLeuGlyIleGlnPro	ThrLeuGlyProProAsnGlnProPro	720	
Db	2209	GACACACGCTACAGTTTCTGGGGATACAG	CCACACTTGGACCTCTAACCAAGCCCT	2268	
RESULT 6	AF291820	3325 bp	mRNA	linear	PRI 15-JUN-2007
LOCUS	AF291820				
DEFINITION					
ACCESSION	AF291820				
VERSION	AF291820.1				
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					

PUBMED	10969042
REFERENCE	2 (bases 1 to 3325)
AUTHORS	Donoghue, M., Woolf, B., Robison, K. and Acton, S.
TITLE	Direct Submission
JOURNAL	Submitted (01-AUG-2000) Cardiovascular Biology, Millennium Pharmaceuticals, Inc, 75 Sidney Street, Cambridge, MA 02139, USA
FEATURES	Location/Qualifiers
source	1..3325
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="X"
	21..2438
CDS	/codon_start=1
	/product="ACE-related carboxypeptidase ACE2"
	/protein_id="AAF9721.1"
	/db_xref="GI:9802433"
	/translation="MSSSSMLLSLVAVTAOSTIEBQAKFLDKFNHAEADLFYQSS
	LSAWNNITREENNOMNNAAGDKSAFLKEOSTLQOMPLQIEONLTVKQLQALQALQ
	NSSVLSDEKSKRLNTILNTSTVYLGKNGCNDPQELLPEGLNETMANSIDYNE
	RLMAWMSMSEYGEKQRLPLIETRYVLEKRNARNHNEEDYGDYWRDYNQVDYDYS
	RQQLDEVEHTFEETKPLVEHLIHAVYAKLMANPSTISPTICLPAHLIDGMKGRFTY
	NIYSLVPEGQKPNIDYVDAMVDQMAOQIFEEAKRFVYSQYPMKTOGEMNSMLT
	DNQVQKACHPADMLDGKDFRILMCTKQITMDPFLAHNGHGLQYDMAVAQEPFL
	RNANGAFHEAAGEIMSLSATPKHLKSLIPEDQENETIEFLKQALITVGTI
	PEYMAEKRRMAYKEGELPKDQMKKMKKEKREIVGVEPDEYICDPSLFTVNS
	DSFTIYTRITRIYQFQROBQALCOAKKKEGLCDLINSDEGOKLFNLRLGKSEPL
	TLAEIVVAGKNNVRLPLNTPEPLFTYIKDQKNNSVCSMDSPYAAOSIKVARS
	KSLADGRAEEMNDNEMYLFRSSVAYAMKQIFELKQNMILFEGEDYRANLKPRI
	FETVAKRNVSDIILPTREVEKAIKMSRSRINDARFLNDNSLEFLGIDPLGPNQPV
	ITVIVGVAGVIVVGIYVILFTGIRDKRKKKNRSGENPYASIDISKENNPQFOM
	DDVQVSE"
BASE COUNT	1022 a 637 c 747 g 919 t
ORIGIN	
Alignment Scores:	
Pred. No.:	0 Length: 3325
Score:	3869.00 Matches: 720
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	9 Gaps: 0
US-09-978-385-2_COPY_19_738 (1-720) x AF291820 (1-3325)	
QY	1 SerThrIleGluGluGlnAlaIleAlaIleThrPheLeuAspIlePheAsnHisGluAlaGluAsp 20
DB	75 TCCACACATGAGGAGACAGGCCAGACATTTTGGACAAGTTTAAACCCAGAAAGCCGGAAGAC 134
QY	21 LeuPheTyrGlnSerSerLeuAlaAspTyrPheAsnTyrAsnThrAsnIleThrGluGluAsn 40
DB	135 CTGTTCTATCAAGATTCACCTTGCTCTTGGAATTTAAACCCAAATATTATCGAAGAGCAAT 194
QY	41 ValGlnAsnMetLysAsnAlaGlyAspIleTyrPheSerAlaPheLeuIleGluGlnSerThr 60
DB	195 GTCCAAAACAAGATTAATGCTGGGACAAATGCTGCTTTTAAAGAACACATCCACA 254
QY	61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
DB	255 CTTGGCCAAAGTATCCACTACCAAGAAATCAAGAAATCTCACGTCACCTCAG 314
QY	81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerIleArgLeuAsnThr 100
DB	315 GCTCTTCAGCAAAAGGGCTTCAGGCTCTTCGAAAGACAAAGACCAACGGTTGACACA 374
QY	101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
DB	375 ATTCTTAATACAAATAGACACATCTACAGTACGAAAGTTGTATACCCAGATTAATCCA 434
QY	121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
DB	435 CAAGAACTGCTATCTACTTGAACACGCTTTGAATGAATAATGGCAACAGTTTGACATAC 499
QY	141 AsnGluArgLeuThrPheAlaIleTyrGluSerTyrPheArgSerGluValGlyLysGlnLeuArgPro 160

```

Db      495 AATGAGAGGCTCTGGGCTTGGGAAAGCTGGAGATCTGGCTGGCAAGCAGCTGAGGCCA 554
Qy      161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
Db      555 TTATATGAAAGATATGAGTGTCTTAAATAATAGATGGCAGAGCAAAATCATATAGAGGAC 614
Qy      181 TyrGlyAspTyrTrrPdaGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
Db      615 TATGGGGAATTAATGAGAGGAGACTAATGAATTAATGAGGATGAGTGTGACTATGACTACAGC 674
Qy      201 ArgGlyGluLeuLeuGluAspValGluHisThrPheGluGluLeuLysProLeuTyrGlu 220
Db      675 CGCGGCCAGTGTGATGAAGATGTGAACATACCTTTGAAGAGATTTAAACCATTAATAGAA 734
Qy      221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
Db      735 CATCTCATGCTCATGTGAGAGGCAAAAGTTGATGATCCATTCCTTCATATCAGCCCA 794
Qy      241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrrPheAsnLeu 260
Db      795 ATTAGATGCCCTCCCTGCTCATTTGCTGTGATATGTG3GGTAGATTTTGGCAAAATCTG 854
Qy      261 TyrSerLeuThrValProPheGlyGluLysProAsnIleAspValThrAspAlaMetVal 280
Db      855 TACTCTTGACAGTTCCTCTTGAGCAGAAACCAACATAGATGTACTGTATGCAATGTG 914
Qy      281 AspGlnAlaTrrPaspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
Db      915 GACCAAGCCCTGGGATGACAGAGAAATATCAAGAGGCCAGAAAGTTCTTGTATCTGTT 974
Qy      301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
Db      975 GGCTTCCTTAATATGACTCAAGATTCCTGGGAAATTCATGCTTAACGGCAGCAAGAAAT 1034
Qy      321 ValGlnLysAlaValCysHisProThrAlaTrrPaspLeuGlyLysGlyAspPheArgIle 340
Db      1035 GTTCAGAAAGCAGCTGCGCATCCACAGCTGGGACCTGGGAGGCCACTTCAGAGATC 1094
Qy      341 LeuMetCysThrTrrPaspAlaThrMetLaspPheLeuThrAlaHisIleGluMetGlyHis 360
Db      1095 CTATATGGCACAAGGTGACAAATGAGACGATTCCTCAGCAGCTCATCATGATGATGGGCAAT 1154
Qy      361 IleGlnTyrAspMetAlaTrrAlaAlaGlnProPheLeuLeuArgAsnGluLysAsnGlu 380
Db      1155 ATCCAGATGATATGCGATATGCTGACCAACCTTTCTCTTAAGAATGAGCTAATATGAA 1214
Qy      381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
Db      1215 GGATTCATGAACCTGCTGGGAAATCATGTCATTTGACAGCCACACTAAGCATTTTA 1274
Qy      401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPro 420
Db      1275 AATTCATATGCTCTTGTGACCCGATTTTCAAGAGCAAAAGAAACAAATAAATTAACCTTC 1334
Qy      421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlyLys 440
Db      1335 CTGCTCAAAACAAGCACTCAGATTTGTTGGAGCTGCGCATTTACTTACATGTTAAGAGAG 1394
Qy      441 TrrPdaGTrpMetValPheLysGlyGluIleProLysAspGluTrrMetLysLysTrrP 460
Db      1395 TGGAGGTGGATGCTTTTAAAGGGGAAATTCCTCAAGACCAATGATGAAAGAGTGGTG 1454
Qy      461 GluMetLysArgGluIleValGlyValValGluProValProHisAspGluThrTyrLys 480
Db      1455 GAGATGAAAGCGAGATATGTTGGGTGGTGGAACTGTGGCCCATGATGAACAATACATG 1514
Qy      481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrrPheArg 500
Db      1515 GACCCCGCATCTCTCTTCATGTTCTAATGATGATCATCATTCATTCGATATTAACAAGG 1574
Qy      501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520

```

```

Db      1575 ACCCTTACCAATTCAGTTTCAAGAAGCACTTTGCACAGCAGCTTAAACATGAGCCCT 1634
Qy      521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
Db      1635 CTGCACAAATGATACATCTCAAACTCATACAGAAAGCTGACAGAAACTGTTCAATATGCTG 1694
Qy      541 ArgLeuGlyLysSerGluProTrrPheLeuAlaLeuGluAsnValValGlyAlaLysAsn 560
Db      1695 AGCCTTGGAATATCAACACCTTGAGCCCTTACGATTTGGAAATGTTCTAGAGCAAGAAC 1754
Qy      561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrrPheLysAspGln 580
Db      1755 ATGAATGTAAGGCACTGCTCAACCTTTGAGCCCTTTTACCTGGCTGAAAGACAG 1814
Qy      581 AsnLysAsnSerPheValGlyTrrPheThrAspTrrPheSerProTrrAlaAspGlnSerIle 600
Db      1815 AACAAATATCTTTGTTGGAGATGAGTACCGATGAGTCAGTCATATGACAGCAAGACATC 1874
Qy      601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrrPaspAsn 620
Db      1875 AAGTGAAGATTAAGCTTAATATCAGCTCTTGAGAGATTAAGCATATGAAATGAAACGACAT 1934
Qy      621 GluMetTyrLeuPheArgSerSerValAlaTrrAlaMetArgGlnTyrPheLeuLysVal 640
Db      1935 GAAATGATACCTGTTCCGATCATCTGTGTGATATGATGATGAGGACGACTTTTAAAGTA 1994
Qy      641 LysAsnGluMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
Db      1995 AAAATCATATATCTTTTGGGAGAGGATGTCGACAGTCCATTAATTTGAAACCAAGA 2054
Qy      661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleGluProArgThr 680
Db      2055 ATCTCTCTTAATTTCTTTGTCTACTGCTACCTTAATATGCTGCTGATATCTTCTAGAACT 2114
Qy      681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
Db      2115 GAAGTTGAAAGGCAATCAGATGTCGCCGAGCCCTATCAATATGATGCTTCGCTGAT 2174
Qy      701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
Db      2175 GACCAACGCTAGATTTCTGGGGATACAGCCACACTTGGACCTCTCAACACCCCTT 2234

```

RESULT 7
AX047758
LOCUS AX047758 3334 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 1 from Patent WO0070032.
ACCESSION AX047758
VERSION AX047758.1 GI:11876765
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3334)
AUTHORS Piddington, C.S., Petrie, C.R., Shoemaker, K.E. and Bishop, P.D.
TITLE Zacc2: a human metalloenzyme
JOURNAL Patent: WO 0070032-A 1 23-NOV-2000;
 ZymoGenetics, Inc. (US)
FEATURES
source Location/Qualifiers
 1..3334
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 35..2452
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="GI:11876766"
 /translation="MSSSSWLLSLVAVTAOSTIEQAKTFDKNHEADLFQSS
 LAGNNVTNTEENVQNMNNAQKWSAFLEQSTLQWYDIOIQLVYLVALQALQO
 NGSVLSDESKRLNTILNTMST IYSGKVCNDPNOECLLPGNEILANSLDYN
 RLMAWESMRSEVQKOLRPLXEYVYLKNEKARHNEHDYDGRGVYENGVDGYS
 RGQILDEYETFEIKPLYLHAAYVRAKLMNAYPSYISPIGCLPHILGDMGRWT
 NLYSLIVPFQKRNIDVTDAMVDQANDACRIFKAEKFPYSGVGLPMNTOGFENSMKT

DPGNYKAVCHPTAMDIGKDRILMCTKVTMDLFAHHEKHIOYMAVAOPELL
 RKGANECHREAVGEIMSSAATPRKIKSIGLSPFOEDNEPNEIFLKOALITVGL
 PETYMEKRWKMYFGEIPKODWKMKMKREIYGVYEPHPHDTYCDPASLEPVS
 DYSFTRYTFLYOFQFOEALCOAAKHGSPHKDCIINSTEQALFNMKLISEPM
 TLLENVVGAKNNMYRPLINFEPLFTYMKOONKNSFVGWSDWSPVADOSIKRISL
 KALDADKAEENDNEMYLEFRSSVAYAMQYFLKKNOMILGEBEDVRYAKPRISFN
 FVTAIPKNVSDIIPREVEKATIRMSRSRIINAFRLNDSIFELGIOPLEGPNOPVS
 IWLIVGYMGIYVIGIYILFTGIRDRKKKKNSRNPASIDISKEENNPFONT
 DDVOTSP*
 BASE COUNT 1011 a 640 c 754 g 929 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3334
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AK047758 (1-3334)
 QY 1 SerThrIleGluGlnAlaIleuSerThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 89 TCACCATTTGAGGACAGGCGCAAGACATTTTGGACAGTTTAACACAGACCCGAAAGAC 148
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPasnTyrAsnThrAsnIleThrGluGluAsn 40
 DB 149 CTGTTCTATCAAGTTCACTTCTCTTCTTGGAATTAATACACCAATTTTCTGAAGAGAT 208
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrPserAlaPheLeuGluGlnSerThr 60
 DB 209 GTCCAAACATGAAATATGCTGGGACAAATGCTGCTTTTAAAGAACAGCTCCAA 268
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 269 CTGGCCCAAAATGATCCACTACACAAATTCAGAACTTCAGCTCAAGCTTCAGCTGAC 328
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysAlaGluAsnThr 100
 DB 329 GCCTTCAGCAAAATGGGCTCTTCAGTCTCTCAGAAAGACAAAGCGTTGAACACA 388
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyValLysAsnProAspAsnPro 120
 DB 389 ATCTAAATACATGACACCACTACAGTACGAAAGTTGTAAACCCAGATTAATCCA 448
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 449 CAAAGATGCTTATCTTGAACACAGTTGAATGAATATGCAAAACAGTTAGACTAC 508
 QY 141 AsnGluArgLeuTyrPalatrrpGluSerTyrPargSerGluValGlyLysGlnLeuArgPro 160
 DB 509 AATGACAGGCTCTGGCGCTGGGAAAGCTGAGATCTGAGCTGGCAAGCGCTGAGGCCA 568
 QY 161 LeuTyrGlnGluTyrValIleLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 569 TTATATGACAGATATGCTGTGAAATAAGACATGCGACAGACCAATCTTTATGAGAC 628
 QY 181 TyrGlyAspTyrTrrpArgLysAspTyrGluValAsnGluValAspGlyTyrAspTyrSer 200
 DB 629 TATGGGATTAATGAGAGAGACTATGAAGTAAATGGGCTAGCTATGACTACAGC 688
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 DB 689 CGCGCCAGATTGATGAAAGATGCGAATACCTTTGAAGAGATTAACCATTAATATGAA 748
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 749 CATCTTCATGCTTATGAGGCAAAAGTTGATGATGATCCCTTATCATACAGTCCA 808
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrrpGlyArgPheTrrpAsnLeu 260
 DB 809 ATTGAGATGCCCTCCCTGCTCATTTGCTGTGATATGTGGGTAGATTTTGGACAAATCTG 868

QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 869 TACTCTTGGACAGTCTCCCTTTGGACAGAAACCAAACTAATGTTACTGATGCAATGGTG 928
 QY 281 AspGlnAlaTrrpAspAlaGlnAlaArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 929 GACCAAGCGCTGGGTGACAGAGAAATTTCAAGAGAGCCGAAAGTCTTTGTATCTGTG 968
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrrpGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 969 GGTCTTCTTAATATGATCACTCAAGATTTCTGGGAAATTCATGCTTAACGAGCCAGAAAT 1048
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrrpAspLeuGlyLysGlyAspPheArgIle 340
 DB 1049 GTTCGAAAGACAGCTGCGCATCCACAGCTTGGGACCTGGGGAAGGCGCACTTCAGGATC 1108
 QY 341 LeuMetCysThrLysValIleThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHis 360
 DB 1109 CTTATGTGCAACAAAGGTGACATGACGACTTCTGACAGCTCATCATGAGATGGGCAT 1168
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1169 ATCCAGTATGATATGAGCATATGCTGACAACTTTCTGTAAGAAATGGAGCTAATGAA 1228
 QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1229 GGATTCATGAAAGCTGTGGGAAATCATGTCACCTTCTGACGCGCACACCTTAACCAATTA 1288
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluTrrpGluIleAsnPhe 420
 DB 1289 AATATCATTTGGTCTTCTCTACCCGATTTTCAAAAGACATGAAACAGAAATTAATTC 1348
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1349 CTGCTCAAAACAGCACTCAGATTTGGGACTGTGCTCATTTACTTCACTGTTGAGAACG 1408
 QY 441 TrrpArgTrrpMetValPheLysGlyGluIleProLysAspGlnTrrpMetLysLysTrrp 460
 DB 1409 TGGAGTGGATGTGTTTAAAGGGGAAATTTCCCAAGACCGGTGAGTGAATAATGGTGG 1468
 QY 461 GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCys 480
 DB 1469 GAGATGAAGCGAGATATGTTGGGTGGTGAACCTGCGCCATGATGAATAACTACTGT 1528
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrrpArg 500
 DB 1529 GACCCCGCATCTCTGTCATGTTCTAATGATTAATCAATTCATTCATATTAACACAGG 1568
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 DB 1589 ACCCTTTACCAATTTCCAGTTTCAAGACACCTTGTCAAGCAGCTAAACATGAAGGCCCT 1648
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
 DB 1649 CTGCACAAATGTGACATCTCAAACTCTACAGAAAGCTGACAGAAACCTTTCAATAATGCTG 1708
 QY 541 ArgLeuGlyLysSerGluProTrrpThrLeuAlaLeuGluAsnValIleValAlaLysAsn 560
 DB 1709 AGGCTTGGAAATATCGAACCTTGGACCTTGAACCTTGAAGTAAATGTTGTAGGACAAAGAC 1768
 QY 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrrpLeuLysAspGln 580
 DB 1769 ATGAAATTAAGGCACTGCTCAAGTACTTGAAGCCCTTAATTTACTGTGCTGAAGACAG 1828
 QY 581 AsnLysAsnSerPheValGlyTrrpSerThrAspTrrpSerProTyrAlaAspGlnSerIle 600
 DB 1829 AACCAAAATCTTGTGTGGATGAGTACCGCTGAGTCCATATGACAGCAACCAAGCAATC 1888
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrrpAsnAspAsn 620
 DB 1889 AATGTAGATTAAGCTTAATAATCAGCTTGTGAGATTAAGCATTAATGATGAGACAGACAT 1948

QY 621 GUMETLrLeuPheArgSerSerValAlaIryAlaMetArgGlnItyrPheLeuVal 640
 DB 1949 GAAATGTACTCTTCGATCATCTGTGATATGCTATGAGGCGACTACTTTTAAAGTA 2008
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGlnAspValArgValAlaAsnLeuIysProArg 660
 DB 2009 AAAAATCAGATATCTTTTGGGAGGAGATGTGCGATATGCTAAATTTGAACACAGA 2068
 QY 661 IleserPheAsnPhenPheValIthrAlaProIysAsnValSerAspIleIleProArgThr 680
 DB 2069 ATCTCTTAAATTTCTTTGTCACCTACCTAAATATGTCGATATCATCTCTAGAACT 2128
 QY 681 GUVAlGluValAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 DB 2129 GAAGTGGAAAGAGCCATCGATGATGTCGAGCCGATCAATGATGCTTCCGTCTGAT 2188
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2189 GACACAGCCCTAGAGTTCTTGCGGATACAGCCACACTTGACCTCTACACAGCCCT 2248

RESULT 8
 HSM800880 3341 bp mRNA linear PRI 18-FEB-2000
 LOCUS Homo sapiens mRNA; cDNA DKFZp434A014 (from clone DKFZp434A014);
 DEFINITION partial cds.
 ACCESSION AL110224
 VERSION AL110224.1 GI:5817159
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 3341)
 Wambolt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
 Direct Submision
 Submitted (15-AUG-1999) MIPS, Am Klopferstr 18a, D-82152
 Martinsried, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 This clone (DKFZp434A014) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at http://www.mips.biochem.mpg.de/proj/cDNA/
 Location/Qualifiers
 1..3341
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp434A014"
 /tissue_type="testis"
 /clone_lib="434 (synonym: hhes3). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 1..2417
 /gene="DKFZp434A014"
 /l..2417
 /gene="DKFZp434A014"
 /note="strong similarity to peptidyl-dipeptidase A"
 /codon_start=3
 /product="hypothetical protein"
 /protein_id="CAB53682.1"
 /db_xref="GI:5817160"
 /translation="SSSWILSLSLVAVTAHSTIEBOAKTFLEKFNHEADLFYOSL
 ASMYNTNTEENYONMNAAGKMSAFLEOSTLAOMPLADEIOMLYALQLOALQON
 GSSVSEDSKRLNTLNTMTSTISGKVCNPNDOECILLEGQNETMANSLDYNER
 IMAWESMRSEVQKOLRPLTEEVYVAKNEARANHEDYDGRGVYEVNGVDGYYSR
 GQIIEVTEFEIKFVLEHIAIVAKLMAVPSITISGLPLPHLGLDMGRWTN
 LYSLVPEGQKPNIDVTDAVQAMDAQKIFKAEKFPVSGLPNTGQFWNSMLTD
 PGVNAQAVCHPTAMDIGKDPRIKCTKVTMDPFLAHHEMHIDYADVAQAOPFLR
 NGANEGHEAVGEINSLSATPRHLKISGLSPDOEDNETEINFLKQALITVGLP
 FTYMLEKRWMMVFKEIIPKIDQMKMKMKMKREIVGVVEVPHDEYTCOPASLHVSND

YSETRVYRLLXOPQPOEALCOAKHEGRLHKDISNSTRAGOKLENNLRIGSEBPT
 LALENTVGAKNMVRPLNTPEPLFTLTKQKNKSRVGNSTDSPIADOSTIKYRISLK
 SALDRAVEMNDNEMYLEFSSVAAYOLKKNOWILFGEEDVAVANLKPRISENF
 EVTAPKNSVDIIPTEVEKAIRMSRSRINDAFRINDNSLEFGLIOPLGPDPVSI
 WLIVGVVMSVIVGIVILFETGIRDRKKKKRSGENPYASIDISGENNPGFQMTD
 DVQTSF"

polysignal 3266..3271
 polysite 3281
 BASE COUNT 1058 a 630 c 740 g 913 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3341
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x HSM800880 (1-3341)

QY 1 SerThrIleGluGlnAlaIystrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 54 TCACACATTTAGGAAGAGCCAGCAATTTTGGACAAGTTTACACAGAGCGGAAGAC 113
 QY 21 LeuPheTyGlnSerSerLeuAlaSerTrpAsnTyRasnThrAsnIleThrGluLys 40
 DB 114 CTGTTCTATCAAACTTCCTGCTTCTGGAATTTAAACCAATATTTACTGAGAGAAAT 173
 QY 41 ValGlnAsnMetLysAsnAlaGluAspLysTrpSerIlePheLeuLysGlnGlnSerThr 60
 DB 174 GTCCAAACATGAAATATGCTGGGACAAATGGCTGCTTTTAAAGACAGTCCACA 233
 QY 61 LeuAlaGlnMetTyProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 224 CTGGCCAAATGTATCCAGCTACAAAGAAATTCAGATTCACAGTCAAGCTTCAGCTCAG 293
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 294 GCTCTTCAGCAAAATGGCTTTCAGTCTTCAGTCTTCACAAAGACAAAGCAAGCTTCACACA 353
 QY 101 IleLeuAsnThrMetSerThrIleTyRserThrGlyLysValLysAsnProAspAsnPro 120
 DB 354 ATTCTAATACAAAGACACCATCTACAGTACTGAAAAGTTTAAACCAAGATATACCA 413
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyR 140
 DB 414 CAAAGATGCTTATTTACTTGAACCAAGGTTGAATGAATATGCAAAACAGTTTACTGACTAC 473
 QY 141 AsnGluAlaGluLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
 DB 474 AATGAGAGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGGTCGCGCAAGCTGAGGCCA 533
 QY 161 LeuTyRglGluGluTyRValValLeuLysAsnGluMetAlaArgAlaAsnHisTyRgluAsp 180
 DB 534 TTATATGAGAGTATGTGCTTGAANAATAGATGGCAGACGAATTCATTATGAGAGAC 593
 QY 181 TyRglAspTyRTrpArgLysAspTyRgluValAsnGluValAspGlyTyRAspTyRSer 200
 DB 594 TATGGGAGTATATGGAGAGGAGACTATGAAATTAATGGGAGATGCTATGACTACAGC 653
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyRglu 220
 DB 654 CGCGGCCAGTTGATGAAGATGTGGAACATCTTGAAGAGATTAACCATTTATGAGAA 713
 QY 221 HisLeuHisAlaTyRValArgAlaLysLeuMetAsnAlaTyRProSerTyRLeuSerPro 240
 DB 714 CATCTTCTGCTATGTATGAGGCAAGTGTATGAAATGCTTCTCTCTATATAGTCCA 773
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGluLysAspMetTrpGlyArgPheTrpHisLeu 260
 DB 774 ATTGATGCTCCCTGCTCATTTGCTGTGTATGTGAGGATGATTTTGGACAAATCTG 833

QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 834 TACCTTTGACAGAGTCCCTTGGACAGAAACCAACATAGATGTACTGATGCAATGGTG 893
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheGlyGlnValIleGlyPhePheValSerVal 300
 DB 894 GACCGAGCCTGGGATGCCACAGAGATTTCAAGGGCCGAGAAAGTTTGTGATCGTT 953
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGlyAsn 320
 DB 954 GGTCTCCCAATATGACTCAAGAGATTCGGGAAATTCATGCTCAAGGACCCAGGAAT 1013
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 DB 1014 GTTCGAAAGACAGCTGCCATCCACAGCTTGGGACCTGGGAGAGGCGCACTTCAGATC 1073
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisIleGlnMetGlyHis 360
 DB 1074 CTTATGTGCACAAAGGTGACATGGACACTTCCGACAGCTCATCATGAGATGGGCAT 1133
 QY 361 IleGlnIYrAspMetAlaTrpAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1134 ATCCAGTMTGATGATGCAATGCTGACACACCTTTTCGTAAGAAATGAGCATATGAA 1193
 QY 381 GlyPheHisGlnAlaValGlyGlnIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1194 GGATTCATGAGAGCTGTGGGAAATCATGTCTCTGACGCCACCTTAACCATTTA 1253
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGlnAspAsnGluThrGlnIleAsnPhe 420
 DB 1254 AATATCATTTGGTCTTCTGTCACCGATTTTCAGAGAAACATGAAACGAAATTAACCTC 1313
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLys 440
 DB 1314 CTGCTCAAAACAGCACTCAGCATTTGTTGGACTCGCCACTTACTTACATTTAGAGAG 1373
 QY 441 TrpArgTrpMetValPheLysGlyGlnIleProLysAspGlnIleTrpMetLysLysTrp 460
 DB 1374 TGGAGGTGGATGCTTTTAAAGGGGAAATCCCAAGACAGTGGATGAAAAAGCTGTGG 1433
 QY 461 GluMetLysArgGlnIleValGlyValValGluProValProHisAspGlnThrTyrCys 480
 DB 1434 GAGATGAGAGCGAGATGTTGGGGTGGTGGAACTGTGCCCATGTAAGAACTACTGT 1493
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArg 500
 DB 1494 GACCCCGCATCTGTTCCATGTTCTTAATGATTAATCATTCATTCGATTTACACAAG 1553
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuGlyGlnAlaIleValSHISGluGlyPro 520
 DB 1554 ACCCTTTCACAAATTCACATTTTCAGAAAGCACTTTGTCAAGCAGCTAAACATGAAGCCCT 1613
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyGlnLysLeuPheAsnMetLeu 540
 DB 1614 CTGCACAAATGTGACATCTCAACTCTACAGAACTGACAGAACTGTTCAATATCTG 1673
 QY 541 ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGlnLysValIleGlyAlaLysAsn 560
 DB 1674 AGGCTTGGAAATACGAACCTGAGACCTTACATTTGGAATAATGTTGAGGACAAAGAAC 1733
 QY 561 MetAsnValArgProLeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuLysAspGln 580
 DB 1734 ATGAAATGTAAGGCGACGCTCACTTTCAGCCCTTAATTTACCTGGCGAAAGACAG 1793
 QY 581 AsnLysAsnSerPheValGlyTyrSerThrAspTrpSerProTyrAlaAspGlnSerIle 600
 DB 1794 AACAAAGATTTCTTTGGGATGAGATGACGACTGGAGTCCATATGACACCAAGCATC 1853
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyLysAspLysAlaTyrGluTyrAsnAspAsn 620
 DB 1854 AAAGTAGAGTAAGCCTTAATATCAGCTTGTGAGATTAAGCATATGAAATGAAGACAAAT 1913
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgIleTyrPheLeuLysVal 640

DB 1914 GAAATGTACCTGTCCGATCATCTGTGATATGATAGGACGATCTTTAAAGTA 1973
 QY 641 LysAsnGlnMetIleLeuPheGlyGlnLysAspValArgValAlaAsnLeuLysProArg 660
 DB 1974 AAAATATGATGATCTTTTGGGAGAGAGATGCTCGATGCTTAATTTGAACCAAA 2033
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 DB 2034 ATCTCCTTAATTTCTTTGTGCACTGCACCCAAAATGTCTGATATCATTTCTAGACT 2093
 QY 681 GluValGlnLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 DB 2094 GAAATGTAAGGCGCATCAGAGATGTCGCCGAGCCGATCAATGATCTTCCGTCGAT 2153
 QY 701 AspAsnSerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2154 GACAAAGCCTTAGAGTTTCTGGGGATACAGCCAAACCTTGACCTTCCTTAACAGCCCT 2213
 RESULT 9
 ARI35177
 LOCUS ARI35177 3396 bp DNA linear PAT 16-MAY-2001
 DEFINITION Sequence 1 from patent US 6194556.
 ACCESSION ARI35177
 VERSION ARI35177.1 GI:14124082
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 3396)
 AUTHORS Action, S. Laurene, and Robison, K. Earl.
 TITLE Angiotensin converting enzyme homolog and therapeutic and diagnostic uses therefor
 JOURNAL Patent: US 6194556-A 1 27-FEB-2001;
 FEATURES
 source location/Qualifiers
 BASE COUNT 1034 a 659 c 772 g 931 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3396
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-978-385-2-COPY_19_738 (1-720) x ARI35177 (1-3396)
 QY 1 SerThrIleGlnGlnGlnAlaLysThrPheLeuAspLysPheAsnHisGlnAlaGlnAsp 20
 DB 136 TCCACATTTGAGGAAAGGCCAAGACATTTTGGACAGATTAAACACGAAAGCCGAAAGC 195
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrAsnTyrAsnIleThrGlnLysAsn 40
 DB 196 CTGTCTATCAAAAGTTCACCTGCTCTGGAATTAACACCAATATTAACGAAAGAT 255
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnIleSerThr 60
 DB 256 GTCCAAACATGAAATATCTGCGGCAAAATGCTGCTTTTAAAGAAACAGTCCACA 315
 QY 61 LeuAlaGlnMetTyrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnIleGln 80
 DB 316 CTGGCCCAATATGATCATCTACAGAAATTCGAATTCACAGTCAAGCTTCAGCTCAG 375
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 376 GCCTTCACCAAAATGGCTTCAGTGTGTCAGAAAGCAAGACGAAACGGTTGAACACA 435
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 DB 436 ATTTCAATATCAATGAGCAGCATCTACAGTACTGGAAGAAAGTTTGAACCCAGATATCA 495

QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluLeuMetAlaAsnSerLeuAspTyr 140
 DB 496 CAGAAAGCTTACTTACTGAAACAGGTTGAATGAATAATGCAAAACAGTTAGACTAC 555
 QY 141 AsnGlnArgLeuThrAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuAlaPro 160
 DB 556 AATGACAGGCTCTGGGCTTGGAAGCTGGAGATGAGAGTGGGCAAGCAGCTGAGGCCA 615
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnIstYrGluAsp 180
 DB 616 TTAATGAAAGAGATATGGTGGCTTGAATAATGAGATGGCAAGACCAATCATTTATGAGGAC 675
 QY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 DB 676 TATGGGATTTATGGAGAGAGACTATGAGTAATGAGGATGATGCTATGCTACAGC 735
 QY 201 ArgGlyGlnLeuIleGluAspValGluIstYrPheGluGluTyrLeuProLeuTyrGlu 220
 DB 736 CCGGCGCAGTTGATTAAGATGTGGAACTACCTTTGAAAGAGATTAAACCATTAATATGAA 795
 QY 221 HisLeuHisAlaTyrValAlaArgAlaLysLeuMetAlaAsnAlaTyrProSerTyrIleSerPro 240
 DB 796 CATCTTCATGCGCATGTGAGGCGCAAGTGTGATGATGCTATGCTTCCATATACAGTCCA 855
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheTrpHisLeu 260
 DB 856 ATTTGATGGCTCCCTGCTCATTTGCTGGTGTGATGATGCTGGGTAGATTGGCAAAATCTG 915
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValIstYrAspAlaMetVal 280
 DB 916 TACTCTTTCAGAGTCCCTTGGACAGAAACCAACATGATCTTACTGATGCAATGAGTG 975
 QY 281 AsnGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 976 GACACAGGCTGGATGTCACAGAAATATTCAGAGAGGCGGAGAACTTTCTGATCTGTT 1035
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 1036 GGGCTTCCATATATGCTCAGAGATTTCTGGAAATTCATGCTAACGGACCCAGGAAAT 1095
 QY 321 ValGlnLysAlaValCysHisSerProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 DB 1096 GTTCAGAAAGACAGTGTGCAATCCACAGCTTGGGAGCTGGGAAAGGCGCACTTACAGATC 1155
 QY 341 LeuMetCysThrLysValIstYrMetAspPheLeuThrAlaHisIstGluMetGlyHis 360
 DB 1156 CTATATGTGCAGAAAGGTGCAATGACACTTCTGACAGCTCATCATGATGAGTGGGCAT 1215
 QY 361 IleGlnTyrAspMetAlaTyrAlaIstYrProPheLeuLeuArgAsnGlyValAsnGlu 380
 DB 1216 ATCCAGTATGATATGCAATATGCTGCAACACTTTTCTGCTAAGAAATGGAGCTATATAA 1275
 QY 381 GlyPheHisGluAlaValGlyIstYrMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1276 GGATTCACATGAGCTGTGGGAAATCATGTCACTTTCGACGCCACACACTAAGCATTTA 1335
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 DB 1336 AATTCATGTTGGTCTTGTGACCCGATTTTCAAGAGACAAATGAAACAGAAATAAATTC 1395
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1396 CTGCTCAAAACAGACACTCAGCATTTTGGGACTCTGCCATTTACTTAAATGTTATAGAG 1455
 QY 441 TrpArgTrpMetValPheLysGlyGluIstYrProLysAspGlnTrpMetLysLysTrpTrp 460
 DB 1456 TGGAGGTGAGTGGCTTTAAAGGGGAAATTCACAAAGACAGATGATGAAAGAGGGTGG 1515
 QY 461 GluMetLysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCys 480
 DB 1516 GAGATGAACCGAGATAGTGGGTGGGAACTGTGCCCATGATGAAACATACTAGT 1575

QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
 DB 1576 GACCCCGCATCTCTGTTCATCTTTCTAATGATTACTCATTTCAATTCATGATTAACAGAG 1635
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 DB 1636 ACCCTTACCAATTCACATTCAGTTTCAGAGACACTTGTTCAGAGCAGCTAAACATGAGGCCCT 1695
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
 DB 1696 CTGCACAAATGTGACATCTCAAACTCTACAGAAAGCTGACAGAAACTGTTCAATATGCTG 1755
 QY 541 ArgLeuGlyLysSerGlnProTrpPheThrLeuAlaLeuGluAsnValIstYrValLysAsn 560
 DB 1756 AGCCTTGGAAATTCAGAACCCCTGGACCTAGCATTTGGAAATGTTGTGAGCAAGAAC 1815
 QY 561 MetAsnValArgProLeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuLysAspGln 580
 DB 1816 ATGAATGTAAAGGCCACTGCTCAACTACTTTCAGCCCTTAATTTACCTGCTGAAAGACAG 1875
 QY 581 AsnLysAsnSerPheValGlyTrpSerThrArgTrpSerProTyrAlaAspGlnSerIle 600
 DB 1876 AACAAAGATTTCTTTTGTGGATGAGTACCGAGCTGAGTCCATATGACAGCAAAAGCATC 1935
 QY 1936 AAAGTGAGATTAAGCTTAATAATCAGCTTGTGAGATTAAGCATATGATGAGACAGACAT 1995
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLysVal 640
 DB 1996 GAAATGTACCTGTTCCGATCATCTGTTCATATGATGATGAGCGAGTACTTTTAAAGTA 2055
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
 DB 2056 AAAAATCAGATGATCTTTTGGGAGAGAGATGTCGATGAGCTAATTTGAAACAGAGA 2115
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 DB 2116 ATCTCTTAAATTTCTTTGCTGACAGCTCACTAAATGATGCTGATATCATTTCCATGAACT 2175
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 DB 2176 GAGCTGAAAGGCGCATGAGATGTCGGGAGCGGTATCAATGATGCTTCCGCTCAAT 2235
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2236 GACAACACCTCAGAGTTTCTGGGATACAGCAACACTTGGACCTCTTAACAGACCCCT 2295
 REFSEQ 10
 LOCUS AX418982
 DEFINITION Sequence 1 from Patent WO0212471.
 ACCESSION AX418982
 VERSION AX418982.1 GI:21523781
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Acton, S., Robison, K.E. and Hsieh, F.Y.
 TITLE Angiotensin converting enzyme homolog and uses therefor
 JOURNAL Patent: WO 0212471-A 1 14-FEB-2002;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source 1..3396
 CDS 82..2499
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD35458.1"
 /db_xref="GI:21523782"

/translation="MSSSMLLSLVAVTAOSTIEBOAKFIDKFNHBAELPFQSS
 LASMNNTNTEENONNNAGDKMSAFILKBOSTLQAWLOEIONLKVLOLOLOO
 KSSVISEDKSKRINTIILNTASTYSTGKVCNPDPOCLLEPGLINTMANSIDYNE
 RLMWMSWSEVKORLPIEETVVLKNNMAYNTEEDGYWRDDEYNGVDGDIS
 NQLEJEDVEHEEIKPLEYELHAYRAALMAYPSYSPICDLVAHLHGDHREMT
 NLYSLVPPGOKPNDIDVADWQADMDORIFKEAEKFEVSGVSLNMGEMENMLT
 DGNVOKAVCHPTAMDLDGDFRILMCTVMDLVAHNEHGHIOYMAVAOPELL
 RANGANGFEHAYGEIWSLSAATPKHLKISGLSPDOENETEINELKQALTYGT
 PPTVMEKRMVAFGEIPIKDOMMKWMEKREIYGVPPHDETCDDPASTLPHVS
 DYSEITRTYRTLYKQFOPALCOAKHBEPLKCDISNTEAGOKLENNLRIGKSEPN
 TLALENNVGAKNMNVAPLNTFEPLETLKDNKNSFWSMSTQPRADQSKVATSL
 KSAIDKAYEMNDENMYLFRSSVAYAMROYFLKVNQMILLFESEYRANLKRISFN
 FEVTAARKNSDIIPTREVEKAIKMSRSRINDAFRLNDSLEBGLIOPATGPNDOPS
 IMLVGVGVAVGVIVGIVILITFGIRDKRKKKNSGENPYASIDISKENNPGFQNT
 DVOQTSF"

BASE COUNT 1034 a 659 c 772 g 931 t

ORIGIN

Alignment Scores:

Pred. No.:	Length:	3396
Score:	Matches:	720
Percent Similarity:	Conservative:	0
Best Local Similarity:	Mismatches:	0
Query Match:	Indels:	0
DB:	Gaps:	0

US-09-978-385-2_COPY_19_738 (1-720) x AX418982 (1-3396)

QY 1 SerThrIleGluLuglnAlaLysThrPheLeuAspLysPheAsnIleGluIleAsp 20

DB 136 TCCACCATTTGAGAACGCCAAGACATTTTGGACAAAGTTTACCAACGACCGCAAGAC 195

QY 21 LeuPheTyrIleSerSerLeuAlaSerTyrPheAsnThrAsnIleThrGluIleAsn 40

DB 196 CTGTCTATCAAAAGTCTACTGCTCTCTGGAATTAACACCAATATTAATCTGAAGAAAT 255

QY 41 ValGlnAsnMetAsnAlaGlyAspLysTyrSerAlaPheLeuGluIleAsnIle 60

DB 256 GTCCAAACATGATATATGCTGGGACAAATGGCTCTCTTTTAAAGGAACGTCACCA 315

QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeu 80

DB 316 CTTCGCCAAATGATCCACTACCAAGAAATTCAGATCTCAGTCACAGTCACAGTCGACG 375

QY 81 AlaLeuGlnIleAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100

DB 376 GCTCTCAGCAAAATGGCTCTGCTGCTGCTCAGAACAAAGACAAACGTTGAAACACA 435

QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120

DB 436 ATTCTAAATACATGACACCATCTACAGTACTGGAAGAAATTTGTAAACCCAGTAAATCCA 495

QY 121 GlnIleCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140

DB 496 CAACAATAGCTTATTACTTGAACCAAGTTTGAATGAATGAATGAACCAAGTTTAACTAC 555

QY 141 AsnGlnArgLeuThrAlaTyrPheLysTyrSerLysValGlyLysGlnLeuAspPro 160

DB 556 AATAGAGGCTCTGGGCTGGGAAAGCTGAGATCTAGGCTGGCAAGCAGCTGAGGCCA 615

QY 161 LeuTyrGluGluTyrValIleLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180

DB 616 TTATATGAGAGTATGGTCTGTAATAATGAGATGCGCAAGAAATCTATTATGAGGAC 675

QY 181 TyrGlnAspTyrTyrPheGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200

DB 676 TATGGGAGTATGTGAGAGGAGACTATGAAGTAAATGGGAGATGCTATGACTACAGC 735

QY 201 ArgGlnLeuIleGluAspValGlnHisThrPheGluGluIleLysProLeuTyrGlu 220

DB 736 CGCGCCAGTGTGAGATGTGGAACATACCTTTGAAGAGATTAACCATTAATATGAA 795

QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240

DB 796 CATCTTATGCGCTATGAGAGGCAAGTGTGATGCCATCTTCTATACAGCCCA 855

QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetProGlyArgPheThrAsnLeu 260

DB 856 ATGGAGAGCCCTCCGCTCATCTTCTGGGATATGAGGGGTACATTTTGGACAAATCTG 915

QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280

DB 916 TACTCTTTGACAGTTCCTTTGGACAGAAACCAACATAGATGTACTGATGCAATGCTG 975

QY 281 AspGlnAlaTyrPheAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300

DB 976 GACCGAGGCTGGAGTGCACAGAGATATATCAAGAGGCCGGAAGTCTTGTATCTGTT 1035

QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrPheLysAsnSerMetLeuThrAspProGlyLys 320

DB 1036 GGTCTCTTATATGATACCTCAAGATTTCTGGGAAATTTCCATGCAAGGACCGCAAGAT 1095

QY 321 ValGlnLysAlaValCysHisProThrAlaTyrPheLeuGlyLysGlyAspPheArgIle 340

DB 1096 GTTCAGAAAGCAGTCTGCCATCCACAGCTTGGGACCTGGGGAAGGCGACTTCAGGATC 1155

QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisIleGluMetGlyHis 360

DB 1156 CTATATGTCACAAAGTGCACAAATGAGCAGCTTCTGACAGCTCATGAGATGGGCAT 1215

QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuAsnGlyValAsnGlu 380

DB 1216 ATCCAGTATGATATGATGATGATGCTGCACACACTTCTTGTGAAGAAATGAGCTAATGAA 1275

QY 381 GlyPheHisGluValAlaGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400

DB 1276 GGAATTCATGAAAGCTGGTGGGAAATGTCATCTTCTGAGGCACACCTTACCAATTTA 1335

QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnLysAsnGluThrGluIleAsnPhe 420

DB 1336 AATCCATTTGCTCTCTGTCACCCGATTTTCAAGAGCATGAAACGAATTAACGTC 1395

QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440

DB 1396 CTGCTCAAAACAGACATCAGATGTTGTCGACTGCGCATTTACTTCAATGTTGAGAGAG 1455

QY 441 TrpArgTyrMetValPheLysGlyGluIleProLysAspArgIleTyrMetLysTyrTrp 460

DB 1456 TGGAGGTGATGCTTTTAAAGGCAATTTCCCAAGACGAGTGAATAAATGCGTGG 1515

QY 461 GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCys 480

DB 1516 GAGATGAACGAGAGATAGTGGGTGGGAACTGTGCCCATGATGAACATATACGT 1575

QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500

DB 1576 GACCCGCAATCTGCTCTCATGTTCTAATGTTACTATTCATTCGATTAATTAACAAGG 1635

QY 501 ThrLeuTyrGlnPheGlnPheGlnAlaLeuCysGlnAlaAlaLysHisGluIlePro 520

DB 1636 ACCCTTACCAATTCACAGTTTCAAGAAGCACTTGTCAAGCAGCAATTAACATGAAGCCCT 1695

QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540

DB 1696 CTGCACAAATGTCACATCTCAAACTTACAGAGCTGACAGAAACTGTCAATATGCTGT 1755

QY 541 ArgLeuGlyLysSerGluProThrPheIleAlaGluIleAsnValAlaGlyValLysAsn 560

DB 1756 AGGCTTGGAAATTCACAAACCTTGACCTTGAAGTGTGTGAGGACCAAGAAC 1815

QY 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrLeuLysAspGln 580

DB 1816 ATGAAATGTAAGGCACTGCTCAACTACTTGTAGCCCTTATTAATGCTGCTGAACACAG 1875

QY 581 AsnLysAsnSerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIle 600

Db 1876 AACAGAAATCTTTTGGGATGAGTACCGACGATGATCCATATGACACCAAGCATC 1935
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPheAsnAsn 620
 Db 1936 AAAGTGAGGATTAACCCCTTAATGCTCTGGAGATTAACCATATGATGATGAAGACAAAT 1995
 QY 621 GluMetTyrIleuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuVal 640
 Db 1996 GAAATGACCTGTTCCATCATCTGTCGATATGCTATGAGCGAGTACTTTTAAAGTA 2055
 QY 641 LysAsnGlnMetIleLeuPheGlyGluLysValAlaTyrAlaLysLeuLysProArg 660
 Db 2056 AAAATACATGATCTCTTTTGGAGAGAGATGCGAGTGGCTAATTTGAACCAAGA 2115
 QY 661 IleSerPheAsnPhenPheValIrrAlaProLysAsnValSerAspIleIleProArgThr 680
 Db 2116 ATCTCTTAAATTTCTTGTCTACTGCGACCTAAAGTGTCTATATCATCTTAAAGACT 2175
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 Db 2176 GAAGTGAAGAGCCATCAGGATGTCGCGAGCCGATCATGATGCTTCCGCTGGAAT 2235
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2236 GACACAGCGCTTGTGCGGAGATACAGCCACACTTGACCTCTTAACCAAGCCCTT 2295
 RESULT 11
 AX431513 3396 bp DNA linear PAT 28-JUN-2002
 LOCUS AX431513 Sequence 1 from Patent WO0239997.
 DEFINITION AX431513
 ACCESSION AX431513
 VERSION AX431513.1 GI:21656345
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS 1
 TITLE Gould, A.E., Stricker-Krongrad, A., Acton, S.L., Brown, J.A., Guan, B.,
 Dales, N.A., Kedambi, V.J., Ocalin, T.D., Patane, M., and Solomon, M.,
 Ace-2 modulating compounds and methods of use thereof
 JOURNAL Patent: WO 0239997-A 1 23-MAY-2002;
 MILENNIUM PHARM INC (US)
 FEATURES
 source Location/Qualifiers
 1..3396
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 82..2499
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD37288.1"
 /db_xref="GI:21656345"
 CDS
 translation="MSSSKLLSLVAVTAOSTIEQAKTFLDKFNHEADLFYQSS
 LASMTNTNITEENVQNNMNAQDKMSAFLEKQSTLAQYPLQELQNLVKLQALQO
 NGSSVLSDEKSKRLNTLNTSTYTGKCNPNPNOCLLEGLNLEIMANSIDYNE
 RLMAFMSRSEVQKOLRLIEYEVVLKNEMARHNVIEGYGVMDVENVDDYDS
 RGQLEDEVEHTEFEELKPLYLEHLAYVAKLNNAYPSYSGICPAHLIGMKMRMT
 NLVSLVFEFGKPNIDYTDADVADQAMDQRIEKAERKFTSVGICPNMTQGVENSMIT
 DPGNVOKVCHPTAMDGKGFRLTCKTAVTMDPLTAHHEMGIQYDMAVAAPFL
 RNCANEGRHAEVAGELMSLAATPKHLKSLIGLSLSPQEDNTEINFLKQALITVGL
 PFTTLEKWRMMAVREKGETPKQDMKKWEMKREIVGVEVPHDETCDPASTLFEVSN
 DYSLTRYTTRLVQFOELQCAAKHGGPLHKCDISNSTEAGOKLFLNMLRGKSEYM
 TLAEVNVGAKNNMVRPLNFLEPLFTLTKQONKNSPYGMSDMSPYADDSIKRIST
 KSALEGDKAYEWNDNEMLEFSSVAYAMROYFLKYKNOMIILEGEDVYANIKRISIN
 FEVTAPKRVSDIIPRETEKAIKMSRIRINDAPFLNDSLEFLDQPLTGPNPPVS
 IMLIVGVMKVIYVGIIVLITFTGIRKRNKKNKARSGENPASTIDISGENNPGFQNT
 DDVQTSF"
 BASE COUNT 1034 a 659 c 772 g 931 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3396
 Score: 3869.00 Matches: 720

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AX431513 (1-3396)
 QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHISGluAlaLys 20
 Db 136 TCCACCATGAGAGAACAGCCCAAGACATTTTGGACAAGTTTAACACGAAACCGAAGAC 195
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerThrPheAsnTyrAsnIleThrGluLysAsn 40
 Db 196 CTTCTTATCAAGTTCACCTCTCTTGGAAATTTAACCAATATTACTGAAGAGAAAT 255
 QY 41 ValGlnAsnMetAsnAlaGluLysAspLysTSPSerAlaPheLeuGlyGlnSerThr 60
 Db 256 GTCCAAATCATGAATAATGCTGGGACAAATGCTGCTTTTAAAGAACAGTCCACA 315
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 Db 316 CTTGCCCAATATATCACTACACACAAATTCAGATTCACAGTCAACCTTCAGCTGAG 375
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerTyrArgLeuAsnThr 100
 Db 376 GCTCTTCAGCAAAATGGGCTTCAGTCTGTCAGAGCAAGCAAGCAAGGTTGAACACA 435
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnPro 120
 Db 436 ATTCTAAATACAAATAGCACCACATCATCAAGTCTGGAAGAAAGTTTAAACCCAGATAACCA 495
 QY 121 GlnGluGlyLeuLeuLeuGluProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyr 140
 Db 496 CAAAGATCTTATTACTTGAACCAAGGTTGATTAATTAAGCAACAGTTTGAAGTAC 555
 QY 141 AsnGluArgLeuThrPalatPglUserTyrArgSerGluValGlyLysGlnLeuArgPro 160
 Db 556 AATGAGAGGCTCTGCGGCTTGGGAAAGCTGGAGATCTAGGTCGCGCAAGCAGTGAAGCCA 615
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHISArgLysAsp 180
 Db 616 TTATATGAAGAGTATGAGTCTTGAAGAAAGAGATGCGCAAGCAAGCAATCATTTAGAGAC 675
 QY 181 TyrGlyAspTyrTyrPargGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 Db 676 TATGGGATTAATGGAAGAGGAGACTATGAATTAATGGGAGAGTGTGACTGATCAGC 735
 QY 201 ArgGlyGlnLeuIleGluAspValGluHISThrPheGluGluIleLysProLeuTyrGlu 220
 Db 736 CGCGGCAAGTTGATGAAGATGTGGAACATACCTTGAAGAGATTAACCATTAATGAA 795
 QY 221 HisLeuHISAlaTyrValAlaArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 Db 796 CATTTATATGCTTATGAGAGGCAAGATTAATGATCCATCTTCCATATCACTGCA 855
 QY 241 IleGlyLysLeuProAlaHISLeuLeuGlyAspMetTyrGlyArgPheTyrPheAsnLeu 260
 Db 856 ATTGAGATGCTCTCCCTGCTATGCTGGGATATGATGGGGTATGATTTGGACAAATCTG 915
 QY 261 TyrSerLeuThrValAlaProPheGlyGlnLysProAsnIleAspValIrrAspAlaMetVal 280
 Db 916 TACCTTTCAGAGTCTCTTTCGACAGAAACCAACATAGATGTTACTATCAACAATGAG 975
 QY 281 AspGlnAlaTyrAspAlaGlnAlaGlyIlePheLysGlnAlaGluLysPhePheValSerVal 300
 Db 976 GACACAGCCGCGGATGACAGAGATATTAAGAGAGCCGACAGAGTTCTTGTATCTGTT 1035
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrGluAsnSerMetLeuThrAspProGlyAsn 320
 Db 1036 GGCTCTTCATATATGACTCAAGGATTCCTGGAAATTCATGCTAACGAGCCAGGAAT 1095
 QY 321 ValGluLysAlaValCysHISProThrAlaTyrAspLeuGlyLysGlyLysPheArgTyr 340

Db 1096 GTTCGAAGACAGTCTGCCATCCACAGCTTGGGACCTGGGGAAGGGGAGCTTCAGATC 1155
 Qy 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHis 360
 Db 1156 CTTATGTGACAAAGATGACATGACAGCTTCTGACAGCTCATCAAGATGGGGAT 1215
 Qy 361 IleglnIlyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuAsnGlyAlaAsnGlu 380
 Db 1216 ATCCAGATGATGATGGCATATGCTGACAAACCTTTTGTGCTTAAGAAATGAGGAGCTATGAA 1275
 Qy 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 Db 1276 GGATTCAGATGAGAGCTTGGGGAATCATGTCATCTTGTGACGCCACCTTAAGATTTTA 1335
 Qy 401 LysSerIleglIleuLeuSerProAspPheGlnIlyrAspAsnGluThrGluIleAsnHe 420
 Db 1336 AAATCCATTTGCTTCTGTCACCCGATTTTCAAGAAAGCAAGAAACGAAATTAACCTTC 1395
 Qy 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 Db 1396 CTGCTCAAAACAGACCTCAGATTTGTGGAGCTGCGCATTAATTACATGTTAGAGAAG 1455
 Qy 441 TrrArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysTrpTrp 460
 Db 1456 TGGAGGTGATGCTTTTAAAGGGGAAATTTCCAAAGACAGCTGATGAAAGTGGTGG 1515
 Qy 461 GluMetLysArgGluIleValGlyValGlyProValProHisAspGluThrTyrCys 480
 Db 1516 GAGATTAACGAGAGATATTGGGGTGGGGAACCTGTGCCCATGATGAACATACATGCT 1575
 Qy 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
 Db 1576 GACCCGCACTCTGTCATGTTCTAATGATTCATCATTCGATTAACAGACAGAG 1635
 Qy 501 ThrLeuTyrGluPheGlnPheGlnIleValAlaLeuCysGlnAlaAlaHisGluGlyPro 520
 Db 1636 ACCCTTTACCAATTTCCAGTTTCAAGAGACCTTTGCAAGCAGCTTAACATGAAGCCCT 1695
 Qy 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyLysLeuPheAsnMetLeu 540
 Db 1696 CTGCACAAATGATGACATCTCAAACTCTACAGAAAGCTGACAGAAACTGTTCAATATGCG 1755
 Qy 541 ArgLeuGlyLysSerGluProTrpThrAlaLeuGluAsnValGlyAlaLysAsn 560
 Db 1756 AGGCTTGGAAATCAAGAACCTGACCTGACCTTGTGAAGATTTTGAAGACAGAAAC 1815
 Qy 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGln 580
 Db 1816 ATGAATGTAAAGGCACCTGCTCACTACTTGAACCTTATTTACTGGCTGAAGACAG 1875
 Qy 581 AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIle 600
 Db 1876 AACCAAGAAATCTTTTGTGGAGTGAAGTACGAGTGGATGATGATGACAGCAACCAACATC 1935
 Qy 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn 620
 Db 1936 AAGGTGGATTAAGCTTAAGCTTAAGCTTGGAGATTAAGCAATGATGAGAGCAACAT 1995
 Qy 621 GluMetLysLeuPheArgSerSerValAlaTyrAlaMetArgLysTrpPheLeuLysVal 640
 Db 1996 GAAATGTACCTGTTCCGATCATCTGTGATGATGATGATGATGATGATGATGATGATGAT 2055
 Qy 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaLysAsnLysProArg 660
 Db 2056 AAAATATCAGATGATTTTGGGAGAGAGATGTCGATGCTTAATTTGAAGCAACAGA 2115
 Qy 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 Db 2116 ATTCCTTTAATTTCTTGTGACCTGACCTAAAGATGTCGTGATATCATTTCTAGAACT 2175
 Qy 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 Db 2176 GAAGTTGAAGAGCCATCAGATGTCCCGAGCCGATATCAATGATGCTTCCGCTGTAAT 2235

Qy 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2236 GACAACAGCCTATAGATTTCGGGATACAGCCAACTGAGACCTCTTAACAGCCCT 2295
 RESULT 12
 AF241254
 LOCUS
 DEFINITION
 Homo sapiens angiotensin converting enzyme-like protein mRNA,
 complete cds.
 ACCESSION
 AF241254.1
 VERSION
 AF241254.1
 KEYWORDS
 GI:8650465
 SOURCE
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 3405)
 Turner,A.J.
 A human homolog of angiotensin-converting enzyme. Cloning and
 functional expression as a captopril-insensitive carboxypeptidase
 J. Biol. Chem. 275 (43), 33238-33243 (2000)
 MEDLINE
 PUBMED
 10924499
 REFERENCE
 2 (bases 1 to 3405)
 Turner,A.J., Hooper,N.M., Hyde,R., Karran,E., Christle,G., Karran,E. and
 Turner,A.J.
 Direct Submision
 Submitted (02-MAR-2000) School of Biochemistry and Molecular
 Biology, University of Leeds, Mount Preston Street, Leeds, West
 Yorkshire LS5 9JT, UK
 FEATURES
 source
 Location/Qualifiers
 1..3405
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="Xp22"
 /tissue-type="Lymphoma"
 104..2521
 /note="Zinc metalloprotease"
 /product="angiotensin converting enzyme-like protein"
 /codon_start=1
 /protein_id="AAF8220.1"
 /db_xref="GI:8650465"
 /translation="MSSSSWLLSLVAVTAOSTIEBOAKTFIDKFNHEADELFQSS
 LAKMTNNTIEENVQNNKNGDKRSAPRKEOSTLAQMYPLEOTONLTIVLOALOO
 NGSSVLSDEKSRILNTITSTYTGKVCNPDPBCLLEGLMEIANSIDYVE
 RLMAWESRSEVGGQLRPLYEYVVLNEMARAHYEDYGVNGDYEVNGVSDTYS
 RGLIEDVEHTEFEIEIKRPLYEHLHAYRAKLNAAVSYISPIGCLPAHLLDGMEFWT
 NLVSLTVPFGKPNIDYDAVDAQMDAQRIFEKAEFEVSVGLPNTQGFWNSMLT
 DPGNVOKAVCHPTAMDLDGKDPRIIMCTKVTMDPFLAHHEMGHIQYDAVAAQPEFL
 RNGAEGHEHVAEGLMSTSAATPKHTKISGLISDPDQEDNTEINFLKQALITVGL
 PFTYLRKRWAVRGEIIPKQNMKKMKEMKREYGVYVEPPEHETVCDAISLEHVN
 DYSLRYRTYRLYOFQFOELCOAKHGGPLHKCDISNTEAGKLNMLRLCSSEW
 TLALENVGAKNMNRPPLNFEPLFTMLKQNNNSFVGSTDSPYADSIKRLI
 KSLADKAYENMDNEMYLEFRSSVAYAMQYFLKYKNOMILGEEEDVAVANLKPISEN
 FEVIAKVNYSIDIPTEVEKAIKMSRSRINKKAPLNNSTLEFLGIOTPLPPNPVS
 IWLIVGYVMGIYVGIYLLFTGIRDRKKKRNKRSNGENPYASIDISKEENPQFQNT
 DDVQSP"
 BASE COUNT 1046 a 651 c 770 g 938 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3405
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AF241254 (1-3405)

LOCUS AB046569 2599 bp mRNA linear PRI 14-APR-2001
 DEFINITION Homo sapiens ace2 mRNA, complete cds.
 ACCESSION AB046569
 VERSION AB046569.1 GI:13516971
 KEYWORDS
 SOURCE Homo sapiens adult cDNA to mRNA, clone_lib:Sugano human cDNA library kata clone:kata4505.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Suzuki,Y., Watanabe,M. and Sugano,S.
 TITLE Cloning, expression analysis and chromosomal localization of a novel ACE like enzyme
 JOURNAL unpublished
 REFERENCE 2 (bases 1 to 2599)
 AUTHORS Komatsu,T., Suzuki,Y. and Sugano,S.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUL-2000) Takami Komatsu, the Institute of Medical Science, Virology; 4-6-1, Minato-Ku, Tokyo 108-8639, Japan (E-mail:komatsu@ims.u-tokyo.ac.jp, Tel:81-3-5449-5283(ex.75283), Fax:81-3-5449-5416)
 FEATURES
 source
 Location/Qualifiers
 1..2599
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="Xp22"
 /clone="kata4505"
 /clone_lib="Sugano human cDNA library kata"
 /dev_stage="adult"
 /country="Japan"
 /note="vector:pME18S FL3, lab host:Top10"
 1..2599
 /gene="ace2"
 55..2472
 /gene="ace2"
 /product="ACE2"
 /protein_id="BAB40370.1"
 /db_xref="GI:13516972"
 /translation="MSSSSWLLSLIVAVTAOSTTEQAKTFDKRNEHAEDEYQSS
 LASWNTNTTEENVONMNAKDGKSAFLKEGSTLAOMPLEIOINLTAKLOALQO
 NGSSVSEDSKRLNTILNTSTITSTGYKCPNDPQECLEPGLNETMANSIDNE
 RLWAMSWRSVEVKOLRPLYEEVVLKNEARAHYEDYGRGDIYVNGVDGIDYS
 RGLDIEVEHTEFEIKPIYEHLYAVRAKLMAAYPSYIPIGLPAHLIGDMGRMT
 NLSLIVPEFGCKNIDVTAMVDQADQRIEKAKEAFVSGVGLPMTCGFENSWLT
 DPENYKAVCHPAMDLKGDERRILACTVTMDPFLAHHEMHIOYDMAVAAOPIL
 RKGANGFHEAVGEIEMSLAAPPKHLKSTGLISPDFOENETETINFLLOALTYGTL
 PFTYMEKRMAYFKGEIRPDQMKRMEKREIYGVVEPVPDHYCDPASLFEHNS
 DYSFIRYRTFLYQFOQELCOAAKHGFLHCDISNSTEAGOKLPNMLRIGKSEPM
 TIALENVYAKNNMVRPLNYEPFLFTWLKDNKNSFVGSTDMSPYADQSLKVRISL
 KSAIDGRAYEMKNNEMVLEFRSSVAVAMROYFLKVNQMLIFGEEDVAVNLKRPISN
 FEVTAPKNVDIIPTEVEKAIKMSRSRIIDAPRLNDSLEFLGIQPTLGPMPQPSV
 IWLIVGVGVGIVGVIVLITGIDRKKKKKAKSGENPVASIDISKGNPNPGQNT
 DDVQTSF"
 BASE COUNT 808 a 503 c 590 g 698 t
 ORIGIN
 Alignment Scores:
 pred. No.: 0 Length: 2599
 Score: 3866.00 Matches: 719
 Percent Similarity: 100.00% Conservative: 1
 Best local Similarity: 99.86% Mismatches: 0
 Query Match: 99.92% Indels: 0
 DB: 9 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AB046569 (1-2599)
 QY 1 SerThrIleGluGluGlnAlaIstThrPheLeuAspIysPheAsnHisGluAlaGluAsp 20
 Db 109 TCCACCATGTGAGAACAGCCAGACATTTTGTGCAAGATTAAACCGAAGCGGAGAC 168

QY 21 LeuPheTyrgInserSerLeuAlaSerTyrPasnTyrAsnThrAsnIleThrgInGluAsn 40
 Db 169 CTGTTTATCAAAATTCACATTCCTTCTTGGAAATATTAACACCAATATTACTGAAGAAAT 228
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspIysTyrSerLalaPheLeuysGluGlnInserThr 60
 Db 229 GTCCAAAACATGAATACGCTGGGGACAAATGGCTCTCTTTTAAAGGAACAGTCCACA 288
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValIlysLeuGlnLeuGln 80
 Db 289 CTTCGCCAAATGATTCACATTCACAGAAATTCAGAAATTCACAGCAAGCTTCAGCTCAG 348
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspIysSerLysAlaGluAsnThr 100
 Db 349 GCTCTTCAGCAAAATGGGTCTTCAGTCTCTCAACAAAGACAAACAGTTGAACTAC 408
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 Db 409 ATTCTAAATACAAATGAGCACCATCTACAGTACTGGAAAAGTTTGTAAACCCAGATATCCA 468
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 Db 469 CAAGAAATCTTATTACTGAACCAAGCTTGAATGAATGAATGCGCAACAGTTTGAAGCTAC 528
 QY 141 AsnGluAlaGluLeuThrPalatTyrGluSerTyrArgSerGluValGlyLysGlnLeuArgPro 160
 Db 529 AATGAGAGCTCTGGCTGGGAAAGCTGGAGATCTAGGCTGGCAAGAGCTGAGGCA 588
 QY 161 LeuTyrGluGluTyrValValIleLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 Db 589 TTATATGAAGAGTATGCTGCTTGAATAATGAGATGGCAAGAGCAATATATGAGGAC 648
 QY 181 TyrGluAspTyrTyrPargIlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 Db 649 TATCGGATTTATTTGAGAGAGAGACTATGAATGAATGGGATGATGCTATGATGATGAC 708
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 Db 709 CGCGCCAGCTGATGATGAAGATGATGACATACCTTTGAAGATTTAAACCTTTATATGAA 768
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 Db 769 CATCTTACAGCTATGATGAGGCAAAAGTTGATGATGCTTCCCTTATATGATGATCA 828
 QY 241 IleGlyCysLeuProAlaHisLeuLeuIlyAspMetTyrPglYaaGpPheThrThrAsnLeu 260
 Db 829 ATTGGATGCCCTCCGCTCATTTGCTGTGATATGATGATGATGATGATGATGATGATGAT 888
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 Db 889 TACTCTTGACAGCTTCCCTTGGACAGAAACCAATAGATGTTACTGATGATGATGATG 948
 QY 281 AspGlnAlaTyrPaspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 Db 949 GACCAAGCCTCGGAGGACAGAGATATTCAGAGGGGCGGAGAGATTTTGTATCTGTT 1008
 QY 301 GlyLeuProAsnMetThrGlnGlyPheThrPglLysAsnSerMetLeuThrAspProGlyAsn 320
 Db 1009 GGCTCTTCTTAATATGATCTCAAGGATTCGGAAATTCATGCTTACAGGACCCAGAAAT 1068
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrPaspLeuGlyLysGlyAspPheArgIle 340
 Db 1069 GTTTCAGAAAGAGCTGCTGACATCCACACCTTGGAGCTGGGAAAGCGGACCTTACAGATC 1128
 QY 341 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisHisGluMetGlyHis 360
 Db 1129 CTATATCTGCAAAAGGTGACAAATGAGACGATCTCTGACAGCTCATCATGATGATGAGGAT 1188
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 Db 1189 ATTCACTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1248
 QY 381 GlyPheHisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400

```

Db      1249 GGAATTCATGAGAGCTTTGGGGAAATCATGTCATCTTTCGACGCCACACTTAAGCATTTA 1308
QY      401 LysSerIleGlyLeuSerProAspPheGlnGluAspGlnThrGluLeuSnphe 420
Db      1309 AATCATGATGCTTCTGTCACCCGATTTTCAAGAGCAAGAAACAAATTAATTAATCTTC 1368
QY      421 LeuLeuGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
Db      1369 CTGGCTCAAAACAGCACTCAGATGTTGGAGCTCTGCCATTTACTTACATGTTAGAGAA 1428
QY      441 TrpArgTrpMetValPheGlyGlyIleProGlyAspGlnTrpMetLysLysTrp 460
Db      1429 TGGAGTGGATGCTCTTTTAAAGGGAAATCCCAAGACCATGGATGAAAAAGTGTGG 1488
QY      461 GluMetLysArgGluIleValGlyValGluProValProHisAspGluThrTyrCys 480
Db      1489 GAGATGAAGCGAGAGATAGTTGGGTGGGAGAACCTGTCGCCCATGATGAACATACCTGT 1548
QY      481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
Db      1549 GACCCCGGACCTCTTCATGTTCTTAATGATTCATCATTCATTCATTCATTCATTCAC 1608
QY      501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaValHisGluGlyPro 520
Db      1609 ACCCTTACCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1668
QY      521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
Db      1669 CTGCACAAATGTCATCTCAAACTCTCAAGAGCTGACAGAACTCTTCAATATGCTG 1728
QY      541 ArgLeuGlyLysSerGluProTrpPheLeuAlaLeuGluAlaValValGlyAlaLysAsn 560
Db      1729 AGCGTTGGAATATCGAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 1788
QY      561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGln 580
Db      1789 ATGATGTAAAGGCACTCTCACTACTTATGACCTTATTTACTTGGCTGAAAGACGAG 1848
QY      581 AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIle 600
Db      1849 AACAGAAATCTTTTGGGAGTAGAGTACCGAGTGGATGATGACAGCAACAAAGCATC 1908
QY      601 LysValArgIleSerLeuLysSerAlaLeuGlyLysPheValTyrGluTrpAsnAspAsn 620
Db      1909 AAAGTAGAGATTAACCTTAATATGACCTTGGAGATGAGCATATGAAATGAGACGACAT 1968
QY      621 GluMetTyrLeuPheAspSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
Db      1969 GAAATGTACCTGTTCCGATCATCTGTTGATATGCTATGAGGCGAGTCTTTTAAAGTA 2028
QY      641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
Db      2029 AAAAATCAGATGATCTTTTGGGAGAGAGATGTGCGAGTGGTAAATTTGAAACCAAGA 2088
QY      661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
Db      2089 ATCTCTTATATCTTCTTGCACATGACCTAAATATGTCATGATTCATTCATTCATTCAT 2148
QY      681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
Db      2149 GAAATGTAAAGGCGCAACAGAGATGTCGCCGAGCGGTATCAATGATGCTTCCCTGCAAT 2208
QY      701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
Db      2209 GACAACAGCTTAGAGTTTCTGGGAGATACAGCCAACTTGGACCTCTTAACAGCCCTT 2268
RESULT 14
AX463938
LOCUS      AX463938      3732 bp      DNA      linear      PAT 16-JUL-2002
DEFINITION      Sequence 71 from Patent WO0140466.
ACCESSION      AX463938
VERSION      AX463938.1      GI:21898986

```

```

KEYWORDS      human.
SOURCE      Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS      Baker, K.P., Bersini, M., Deforge, L., Desnoyers, L., Filvaroff, E.,
Gao, W.O., Gerritsen, M.B., Goddard, A., Godowski, P.J., Gurney, A.L.,
Sherwood, S., Smith, V., Stewart, T.A., Tumas, D., Watanabe, C.K.,
Wood, W.L. and Zhang, Z.
TITLE      Secreted and transmembrane polypeptides and nucleic acids encoding
JOURNAL      Patent: WO 0140466-A 71 07-JUN-2001;
Genentech Inc. (US)
FEATURES
SOURCE      Location/Qualifiers
1..3732
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      1137 a 722 c 821 g 1052 t
ORIGIN
Alignment Scores:
Pred. No.:      1,296-312      Length:      3732
Score:      3720.00      Matches:      717
Percent Similarity:      83.88%      Conservative:      1
Best Local Similarity:      83.76%      Mismatches:      2
Query Match:      96.15%      Indels:      138
DB:      6      gaps:      1
US-09-978-385-2_COPY_19_738 (1-720) x AX463938 (1-3732)
QY      1 SerThrIleGluGlnGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
Db      94 TCCACCTTATAGAGACAGCGCCAGACATTTTGGACAAGTTTAAACACAGAACCGGAAGAC 153
QY      21 LeuPheTyrGlnSerSerLeuAlaSerTyrAsnTyrAsnThrAsnIleThrGluGluAsn 40
Db      154 CTGTTCTATCAAAAGTTCACTTCTGCTTGGAAATTAACCAACCAATTTCTGAAAGAAAT 213
QY      41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
Db      214 GTCCAAACATGAAATATGCTGGGACAAATGCTGCTTTTAAAGAGACAGTCCACA 273
QY      61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
Db      274 CTGGCCAAATGATACCATACAGAAATTCAGAAATCTCACAGTCAAGCTTCAAGCTGCA 333
QY      81 AlaLeuGlnGlnAsnGlySerSerValLysSerGluAspLysSerLysArgLeuAsnThr 100
Db      334 GCTCTTACGAAATATGCTTCCAGTCTCTCAGAAAGACAAAGCAACGAGTTGAACACA 393
QY      101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyValCysAsnProAspAsnPro 120
Db      394 ATCTCAATATACATGACAGCACTCTACATCTGGAAGAGTTGTAACCCAGATATCCA 453
QY      121 GlnGluCysLeuLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
Db      454 CAGAAATGCTTATTTACTTGAACACAGGTTGAATGAATATGCAACGATTAACTACAC 513
QY      141 AsnGluArgLeuThrAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
Db      514 AATGAGAGGCTCTGGGCTTGGGAAAGCTGGAATCTGAGTCTGGCAACAGCTGAGGCCA 573
QY      161 LeuTyrGluGluLysTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
Db      574 TTAATGAAGAGTATGAGTGTCTTGAAGAAATGAGATGCGCAAGCAACATATATGAGGAC 633
QY      181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAlaGlyValAlaGlyLysTyrAspTyrSer 200
Db      634 TATGGGAGTATTTGAGAGAGAGACTATGAAATGAGGATGATGCTATGACTACAGC 693
QY      201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220

```


QY 561 MetasValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGln 580
 |||||||
 DB 1669 ATGAAATGTAGGCCACCGCTCACTACTTGTAGAGCCCTTATTTACCTGCGTGAAGACAG 1728
 |||||||
 QY 581 AsnLysAsnSerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIle 600
 |||||||
 DB 1729 AACAGAAATCTCTTGTGGATGTAGTACGACAGGAGGATGCTCAAGG----- 1773
 |||||||
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyLysPylValTyrGluTyrPheAsnAspAsn 620
 |||||||
 DB 1773 ----- 1773
 |||||||
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
 |||||||
 DB 1774 GAATCTTCAATCTCTTGAATGTGCTCTGCTGCGCAACAGTGAAGATGTTGTT----- 1827
 |||||||
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGlnAspValArgValAlaAsnLeuLysProArg 660
 |||||||
 DB 1828 -----TTGTTCTCTACAGGAGGAGAGATGTGGAGTGGCTTAATTGAAACCAAGA 1878
 |||||||
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 |||||||
 DB 1879 ATCTCTTAAATTTCTTGTCTACTGCACTAAATAATGTGTGATATCATCTCTAGACT 1938
 |||||||
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 |||||||
 DB 1939 GAATGTGAAGAGCCATCAGAGATGTCCCGAGCCGTATCATATGATGTTCCGCTGAAT 1998
 |||||||
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 |||||||
 DB 1999 GACAACAGCCCTAGATGTTGTGGGATACAGCAACACTGTGACCTCTCAACAGCCCT 2058
 |||||||

RESULT 16
 AX047762 2638 bp DNA linear PAT 15-DEC-2000
 LOCUS Sequence 5 from Patent WO0070032.
 DEFINITION AX047762
 ACCESSION AX047762
 VERSION AX047762.1 GI:11876768
 KEYWORDS
 SOURCE
 ORGANISM

Mus sp.
 Mus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2638)
 REFERENCE
 AUTHORS
 Piddington, C.S., Petric, C.R., Shoemaker, K.E. and Bishop, P.D.
 TITLE
 Zage2: a human metalloenzyme
 JOURNAL
 Patent: WO 0070032-A 5 23-NOV-2000;
 ZymoGenetics, Inc. (US)

FEATURES
 source
 Location/Qualifiers
 1..2638
 /organism="Mus sp."
 /db_xref="taxon:10095"
 106..2523
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC18992.1"
 /db_xref="GI:11876769"
 /translation="MSSSSMILLSLVAVTTAASLTENAKTFLNNFQAEADLSYSS
 LASNNYNTTEENAKMSEAKWAFYEEOSKTAQSFSLDEIOTPIIKROALDLO
 SSSALSDKRRKQNTLNTSTYTSYSGKCNPRNPECILLEGDEIMATSDYNS
 RLMMEGRRAEVGKQRLPEYVLLNEMARANNIDYGGYMGDIEAGADSDYN
 RNQIIEVETFAIKRLEHLAYVRKLDYPSYISPGCLPALLDMDGREFV
 NLYLTPVFAQPNIDYDAMNQGMAERIFQAEKFFSPVGLPMLDQFMASMTL
 EPADGRKVVCHPTAMDLGDFRIKMCCTVTMDNPLFAHMEGHIQYDMAYARPEL
 RGNAGNEGHEAVGEMLSAATPKHLSIGLPSDPODSETEINFLKQALITVGL
 PPTYLEKRRMAYRGELPKQMMKKMKMKREIVGVPEPHDETCDASLEHVS
 DYSEIRITRTIYFOEALCQAKNGSLKCDISNTEAGCKLKMSTLSESE
 TKALENVVGARNDKVKPLNTEPLFTWLEQNRNSFVWNTWESPADSIKRVISL
 FSLGANAAYEWNTNEMELFRSSVAYARKFYSIIKNOTVPLEDEVDYDLKPVSY
 FEVSPQNVSDVIRSEVEDAIRMSGRINDVFGINDNSLEFLGIHPTLPPQPPYT
 IMLIFGVNAAVAVGIIILVTGIKKKKKKNETKRENPYDSMDIKGSMNGFQNS
 DDAQTSF"

BASE COUNT 801 a 557 c 611 g 669 t

ORIGIN

Alignment Scores:
 Pred. No.: 3,766-272 Length: 2638
 Score: 3251.00 Matches: 598
 Percent Similarity: 89.86% Conservative: 49
 Best Local Similarity: 83.06% Mismatches: 73
 Query Match: 84.03% Indels: 0
 DB: 6 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AX047762 (1-2638)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHleGluAlaGluAsp 20
 |||
 DB 160 TCCCTCAGCGAGGAAATGCCAAGACATTTTAAACAACATTAAACAGAGGCTGAAGAC 219
 |||
 QY 21 LeuPheTyrGlnSerSerIleuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsn 40
 |||
 DB 220 CTGTCTTATCAAAAGTTCACCTCTCTTGGAAATTATATCAATTAATCTGAAAGAAAT 279
 |||
 QY 41 ValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
 |||
 DB 280 GCCCAAAAGATGATGATGAGCTGCAGCCAAATGCTGCTTTATGAAGAACAGCTTAAG 339
 |||
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 |||
 DB 340 ACTGCAAAAGTTTCTCACTACAGAAATCCAGACTCCGATCATCAACGCTCACTACAG 399
 |||
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 |||
 DB 400 GCCCTTACCAAAAGTGGCTTCTACACCTCTGACGACACGACAGAACAAACAGTTGAACACA 459
 |||
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnPro 120
 |||
 DB 460 ATCTGTAAACACCATGACGACCATTTACACTACTGGAAGAAAGTTGCAACCAAGAACCA 519
 |||
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 |||
 DB 520 CAAGAAATGCTTATTAATTGAGCAGAGATGTGAATGAAATGAGCGACACACATCTAC 579
 |||
 QY 141 AsnGluArgLeuThrAlaTrpGluSerTrpArgSerGluValGlyGlnLeuArgPro 160
 |||
 DB 580 AACTTAAAGCTCTGGGCAAGGAGGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 639
 |||
 QY 161 LeuTyrGluGluTyrValValIleuLysAsnGluMetAlaArgAlaAsnHleTyrGluAsp 180
 |||
 DB 640 TTGTATGAAGAGATGTGCTCTGAAAAACGAGATGGCAAGACAAACATTAATAGCAG 699
 |||
 QY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 |||
 DB 700 TATGGGATTAATTGGAGAGGAGGCTATGAAGCAGAGAGGAGATGGCTACCACTATTAAC 759
 |||
 QY 201 ArgGlyGlnLeuIleGluAspValGluHleThrPheGluGluIleLysProLeuTyrGlu 220
 |||
 DB 760 CATAACCACTGATTAATGAATGTAGACCTACCTTGCAGAAATTAAGCAATGATGATGAG 819
 |||
 QY 221 HisLeuHisAlaTyrValAlaArgAlaLysLeuMetAlaLysProSerTyrIleSerPro 240
 |||
 DB 820 CACTTCAAGGCTATGTAGAGAGAGGATGATGATGATGATGATGATGATGATGATGATG 879
 |||
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheThrAsnLeu 260
 |||
 DB 880 ACTGAGTGGCTCTCTCCATTTGCTTGGTATGATGATGATGATGATGATGATGATGATG 939
 |||
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 |||
 DB 940 TACCCTTGAATCTTCTCTTGTGCAAGAAACCAACATGATGATGATGATGATGATGATG 999
 |||
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 |||
 DB 1000 AATCAGAGGCTGGATGACAGAAAGATATTCAAGAGCAGAGAAATTTCTTCTGCTGT 1059
 |||
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGluAsn 320

```

Db 1060 GGCCTCTCTATGATCAAGATCTGGCAAACTCTGCTGATGCGCAGAGAT 1119
QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgLeu 340
Db 1120 GGGCGGAAAGTTGTCTGCCACCCACAGCTGGGATCTGGGACACGGAGACTTCAGAAATC 1179
QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHis 360
Db 1180 AAGATGTGTCAAAAGGTCAACATGAGCAAACTTCTGACACCCATCAGCATGGGACAC 1239
QY 361 IleGluTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
Db 1240 ATCCATATATGATGATATGCGCAGGCAACCTTCTGTAAGAAACGAGCAATGAA 1299
QY 381 GlyPheHisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
Db 1300 GGGTTCATGACCTGTGGAGAAATCATGCTCTTGCACACTCCCCCAAGCATCTG 1359
QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsn 420
Db 1360 AAATTCATGTTGCTTCTGCGCATCCGATTTCAAGAAATGACGAAACAGATAAACTTC 1419
QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlyHis 440
Db 1420 CTACTGAACAGCATTTGACAAATGTTGGACACTACCTTTACTTACTGATGAGAGAG 1479
QY 441 TrpArgTrpMetValPheLysGlyLysIleProLysAspGlnThrMetLysLysTrp 460
Db 1480 TGGAGGTGATGTTGCTTGGGGTGAATCCCAAGAGACAGATGAGTGAAGATGCTG 1539
QY 461 GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCys 480
Db 1540 GAATATTAAGGGGAGATCGTTGGTGGAGACCTCTGCTCATGATGATAAATCACTACT 1599
QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrpArg 500
Db 1600 GACCTGCAATCTGCTTCCATGTTCTTAATGATTAATCACTTCAATTCATGATTAACCA 1659
QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
Db 1660 ACCATTACCAATTCAGATTCAGAGAGCTCTTTCAGACAGCTAAGATATATGCTCT 1719
QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
Db 1720 CTGCAACAATGTGACATCTCAAAATTCCTGAGGCTGAGGAGAGTGTCTCAAGATGCTG 1779
QY 541 ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsn 560
Db 1780 AGCTCTGAAATTCAGAGCCCTGAGCAACAAAGCTTGGAAATGCTGTAGAGACCAAGAT 1839
QY 561 MetAsnValArgProLeuLeuAsnTyrPheGlnPheLeuPheThrTrpLeuLysAspGln 580
Db 1840 ATGAGATGTAACCACTGCTCAATTAATTCACACGCTGTTGACTGCTGCAAAAGAGAG 1899
QY 581 AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIle 600
Db 1900 AACAGAAATTCCTTTGTGGGTGGAACACTGATGAGAGCCCAATATGCGACCAAGCAT 1959
QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn 620
Db 1960 AAGATGAGATTAAGCTTAATAATCACTCTTGGAGCTAATGATGATGAGACCAACAC 2019
QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetAlaGlnTyrPheLeuLysVal 640
Db 2020 GAAATGTTCTGCTTCGATCACTGTTGATGCAATGAGCAAACTATTTTCAATATATC 2079
QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
Db 2080 AAAAACACAGACAGTTCCTTTCTAGAGAAAGATGACAGAGTGGAGATTTGAACCAAGA 2139
QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680

```

```

Db 2140 GTCTCCTCTACTCTTTGTCACCTACCCCAAAATGATGTCATCTCTGAGAGT 2199
QY 681 GluValGlnLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
Db 2200 GAAGTTGAAGATGATCATGAGATGCTGGGGCCCAATCAATGATGCTTGGCTGAAAT 2259
QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGluProPro 720
Db 2260 GATTAACAGCTGTGAGTCTTCTGGGATTCACCAACACTTGAAGCCACTTACCAAGCTCT 2319

RESULT 17
BC026801
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
FEATURES
source
CDS

```

2739 bp mRNA linear ROD 07-ADG-2002
 Mus musculus, RIKEN cDNA 2010305L05 gene, clone MGC:25940
 IMAGE:4236529, mRNA, complete cds.

house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2739)
 Strausberg, R.
 Direct Submission
 Submitted (04-Apr-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (ILNL)
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contract: amgdbcm.tmc.edu
 Gunaratne, P. H., Garcia, A. M., Lu, X., Hulyk, S. W., Hale, S. M.,
 Yoon, V. S., Kovis, C. R., Lawrence, S., Martin, R. G., Muzny, D. M.,
 Richards, S., Gibbs, R. A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 31 Row: e Column: 4
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, similarity but not identity to protein.

location/Qualifiers
 1..2739
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/A"
 /clone="MGC:25940 IMAGE:4236529"
 /tissue_type="Kidney, normal, 5 month old male mouse."
 /clone_id="NCI_CGAP_Kid14"
 /lab_host="DH10B"
 /note="vector: pCMV-Sport6"
 176..2593
 /codon_start=1
 /product="RIKEN cDNA 2010305L05 gene"
 /protein_id="AAH26801.1"
 /db_xref="GI:20071164"
 /db_xref="LocustID:70008"
 /translation="MSSSWLLISLVAVTTAOSITZENATFLNNFNOAEADISYSS
 LASNRYNTNTEENAQMSAAKWSAFYEOSKTAOSFSLDIOTPIIKRQALAQ
 SSSALSADNKOLNLTNLTNMTSTYSGKVCNPKNDECLLEPGIDEIMASTDYS
 RLWAMEGRAEVRKOLPLNLTNMTSTYSGKVCNPKNDECLLEPGIDEIMASTDYS
 RNOLIEDVETFEAIKPLYLHLAAYVRKIMLDYPSYISPTGLPAHLIGDMGRWT
 NLYLVEFAOKRPIVDYTDAMNMGDAEIRIPEAKRFPYSGVGLPHMTGTFWANSULT
 EPDGRKVVCHPFLAMLDGHDFRIKCKTKYMDNFTLAHHEMGIOTDMATYAKOPFL

BASE COUNT 832 a 577 c 620 g 710 t

ORIGIN

BASE COUNT	832 a	577 c	620 g	710 t
ORIGIN				
Alignment Scores:				
Pred. No.:	3,95e-272			Length: 2739
Score:	3251.00			Matches: 598
Percent Similarity:	89.86%			Conservative: 49
Best Local Similarity:	83.06%			Mismatches: 73
Query Match:	84.03%			Indels: 0
DB:	10			Gaps: 0


```

Qy 301 GlyLeuProAsnMetThrGlnGlyPheThrGluAsnSerMetLeuThrAspProGlyAsn 320
Db 1141 GGCTCTCTCATAGTACATCAAGATTCCTGGCAACATCTATGCTGATGAGCCAGAGAT 1200
Qy 321 ValGlnLysAlaValCysHisProThrAlaThrAspLeuGlySerGlyAspPheArgIle 340
Db 1201 GCGCGGAAAGTGTCTCCACCCACAGCTGGATCTGGACACGAGAGCTTCAACAAATC 1260
Qy 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHis 360
Db 1261 AAGATGTGTACAAAGGTCACAAATGACAACTTCTTGACACCCATCAGAGATGGGACAC 1320
Qy 361 IleGlnIleThrAspMetAlaThrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
Db 1321 ATCCATATATGACATGATATGCGACGACCACTTCTGCTAGAAACGAGCAATGAA 1380
Qy 381 GlyPheHisGluAlaValIleGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
Db 1381 GGGTTCATGAAAGCTGTGGAAATCATGCTACCTTCTGACCTACCCCAAGCATCTG 1440
Qy 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAsnGlnThrGluIleAsnPhe 420
Db 1441 AAATCATGTGCTCTTCCCATCGCATTTTCAAGAAATAGCGAAACAGAGATTAACCTTC 1500
Qy 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyMetLeuGluLys 440
Db 1501 CTACTGAAACAGGATGACAAATGTTGGAACACTACCGTTTACTTACATGTTAGAGAAAG 1560
Qy 441 ThrArgThrMetValPheLysGlyGluIleProLysAspGlnThrMetLysLysThrPro 460
Db 1561 TGGAGGTGATGATGCTTTCGGGGTGAATATCCCAAGAGAGATGAGTGAAGAAAGTGGTG 1620
Qy 461 GluMetLysArgGlnIleValIleValIleProValProHisAspGlnThrTyCys 480
Db 1621 GAGATGAAAGGGGAGATCGTTGCTGTGGAGCCCTGCTCATGATGAACATCACTGT 1680
Qy 481 AspProLysSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
Db 1681 GACCTCATCTCTGTTCCATGTTTCAAGATTAACATTCATTCATATTCACACAAAGG 1740
Qy 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnIleAlaLysHisGluGlyPro 520
Db 1741 ACCATTACCAATTCAGATTTCAGAAAGCTTGTGTACAGAGCTAAGTAAATGTTCT 1800
Qy 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyLysLeuPheAsnMetLeu 540
Db 1801 CTCGACAAATGTGACATCTCAATTCACAGAGCTGGGACGAAAGTTGCTCAAGATCTG 1860
Qy 541 ArgLeuGlyLysSerGluProTyrPheLeuAlaLeuGluAsnValIleGlyAlaLysAsn 560
Db 1861 AGTCTTGGAAATTCAGAGCCCTGACCAAGCCCTTGGAAATGTTGGTAGAGAGCAAGAT 1920
Qy 561 MetAsnValArgProLeuLeuAsnTyrPheGlnProLeuPheThrThrPheLysAspGln 580
Db 1921 ATGATGTGAAACACACCTGCTCAATTAATTCACACCTGTTGTGACTGCTGCAAGAGCAG 1980
Qy 581 AsnLysAsnSerPheValIleTyrPheSerThrAspTyrSerProTyrAlaAspGlnSerIle 600
Db 1981 AACAGAAATCTTGTGGGGTGAGACAGTGAATGAGCCCATATGCGGACCAAGCATT 2040
Qy 601 LysValArgIleSerLeuLysSerAlaLeuGlyLysPylsAlaTyrGluTyrPheAsnAspAsn 620
Db 2041 AAATGTGAGGATTAACCTTAATTCAGCTTGGACCTAATCATATGTAATGAGCAACCAAC 2100
Qy 621 GluMetTyrLeuPheAspSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
Db 2101 GAAATGTTCCTGTTCCGATCATCTGTGCAATAGCCCATGAGAAAGTATTTTTCATATATC 2160
Qy 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
Db 2161 AAAAACAGACAGTCTCTTCTAGAGAGAGATGATGAGTGAATTTGAAACCAAGA 2220
Qy 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680

```

```

Db 2221 GTCCTCTCTACTCTTTGTCACTACCCCAAAATGTCGTGATGATCTCTAGAGT 2280
Qy 681 GluValGlnLysAlaIleArgMetSerArgIleAsnAspAlaPheArgLeuAsn 700
Db 2281 GAAGTTGAAGATGATCCATGATGATGCTGGGGCCCATCAATGATGCTTGGCCGTGAT 2340
Qy 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
Db 2341 GATTAACAGCTGAGTCTTCTGGGATTAACCAACACTTGAAGCACCCTTACAGAGCTCT 2400

RESULT 19
AX047765
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLES
JOURNAL
FEATURES
source
1. 2638
/organism="Mus sp."
/db_xref="taxon:10095"
106. 2523
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:11876772"
/translation="MSSSWLLSLVAVTTRQSLTEENAKFLNNFNOEADLSYSS
LASWYNNTNTEENAKMSEAAKSAVEEOSTKPAOSFSLQETIPIKROLALQ
SSSLASLQKNTLINTMSTIVSTKVNPNPOCLILEGDEIMATSTDYNS
RLAMEGMEAEVKQRLPLEYERYVLRKEMRANRYNDYGDYRDEADYDYN
NROLVEDYRPAETKIPLEYEHVLRKEMRANRYNDYSPGCPAILHIGMDREFM
NLYPLTPPAOKPNDIYDAMNOMKDEIRFOEAREKFFVGVGLPHMGCRWASMTL
EPADGRKYVCTHNDLGHGPRIMCKTQVIMDNFLIRHMEGHIQDMATARPFL
RNGANEGEHAEVGLMSLAATPKHLKIGLPSDFQDSLETINFLKQALTVGL
PTVYLMERKRWAVFGEIPEKEDMKMKWEMREIYGVPELRDETYCDPALSLEHVS
DYSEFRTYRTYFOFOBALCOAKYNGSLHKCDISNTEAGOKLKLMSLGNSEP
TALLENVGARMDYKPLINTYFOPLFDMKKQNRSEFGMTSEMPYDOSIKRISL
KSLGANAVENTNEMFLPRSSVAAMKSSLIKNQVPLLEDYRVSOLKPRVSVY
FVYTSFQNVSDYTPRSEYEDAIRMSRGINDYFGLNDSLEFLGHTPLEPYPPT
IWLIFGVVMAIIVVGIILITVTGIGKRRKNETKRENPYDSMDIGKGSNAGFQNS
DDAQTSF"
BASE COUNT 797 a 557 c 615 g 669 t
ORIGIN
Alignment Scores:
Pred. No.: 1,37e-270 Length: 2638
Score: 3233.00 Matches: 596
Percent Similarity: 89.58% Conservative: 49
Best Local Similarity: 82.78% Mismatches: 75
Query Match: 83.56% Indels: 0
Gaps: 0
US-09-978-385-2_COPY_19_738 (1-720) x AX047765 (1-2638)
Qy 1 setThrIleGlnGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
Db 160 TCCCTCACCGAGAGAAATGCCAAGCATTTTAAACACTTAAACAGAGAGGCTGAACAC 219
Qy 21 LeuPheTyrGlnSerSerLeuAlaSerTyrAsnTyrAsnIleThrGluGluAsn 40
Db 220 CTGTCTTATCAAAAGTCACTCTCTGATTAATTAATCAATTAATCAATTAATTAAT 279
Qy 41 ValGlnAsnMetAsnAlaLysAspLysTyrPheSerAlaPheLeuLysGluGlnSerThr 60

```


Db	280	GGCCAAAAGATGATGAGCGCTGACGCCAAATGGTGCCTTTATATACAAACGCTTAAG	339
Qy	61	LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValIlysLeuGlnLeuGln	80
Db	340	ACTGCCCAAGTTTCTCACTACAGAAACCCAGACTCCCATCATCAACAGGTCATACAG	399
Qy	81	AlaLeuGlnGlnAsnGlySerSerValIleSerGluLysPlySerIysArgLeuAsnThr	100
Db	400	GGCCTTCACCAAGAGTGGCTTCACAGACCTTCAGCAGACAAACAAACAAAGTTGAACACA	459
Qy	101	IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro	120
Db	460	ATTTCGAACCCATGATGGCCACCATTTACAGTACTGGAAAAGTTTGCAACCCAGAAACCCA	519
Qy	121	GlnGluCysLeuLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr	140
Db	520	CAGAATGCTTTATCTTGAGCCAGAGATTGGATGAAATATATGGCAAGCAAGACATAC	579
Qy	141	AsnGluArgLeuThrPalatTrrpGluSerTrrpAspSerGluValGlyLysGlnLeuArgPro	160
Db	580	AACTTACGCTCTGGGCGATGGAGGGCTGGAGCGCTGAGGCTTGGCAAGCAGCTGACGGC	639
Qy	161	LeuTyrGlnGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp	180
Db	640	TTGTATGAAAGATATGTGGCTCTGAAAAACGATGGCAAGGCAACCAATTTATTAACGAC	699
Qy	181	TyrGlyAspTyrTrrpArgGlyAspTrrpGluValAsnGlyValAspGlyTyrAspTrrpSer	200
Db	700	TATGGGAGTTTATTGGAGAGGGACTATGAAAGCAGAGGGAGCGAGATGGCTACAACTATAC	759
Qy	201	ArgGlyGlnLeuIleGlyAspValGlnHisThrPheGlnGluIleLysProLeuTyrGlu	220
Db	760	CGTAACCGATGATGAAGATGTACAAACGTAACCTTCGCAAGATCAACGCAATTTATGTAG	819
Qy	221	HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrTrpLeuPro	240
Db	820	CATCTTACGCTATGTAGGAGGAAAGTTGATGATTAACCTCTCTACTACACGCCCC	879
Qy	241	IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrrpGlyArgPheThrAsnLeu	260
Db	880	ACTGGATCCCTCCCGCCCATTTGGTGGTGTATGTGGGAGACATTTTGGCAAAATCTG	939
Qy	261	TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal	280
Db	940	TACCTTGAATGCTTCCTTTGGCAGCAAGAAACCAACATAGATGTACATGATCAATGTAG	999
Qy	281	AspGlnAlaTrrpAspAlaGlnArgIlePheLeuGluValGluLysPheValSerVal	300
Db	1000	AATCAGGGCTGGGATGCAAGAAAGGATATTCAAGAGGACAGAAATTTGTTGTTCTCTT	1059
Qy	301	GlyLeuProAsnMetThrGlnGlyPheThrGluAsnSerMetLeuThrAspProGlyAsn	320
Db	1060	GGCCTTCTCATATGACTCAAGAGATCTGGGCAAACTCTTATGCTGACACGACCAACAT	1119
Qy	321	ValGlnLysAlaValCysHisProThrAlaTrrpAspLeuGlyLysGlyAspPheArgTle	340
Db	1120	GGCGGAAAGTGTCTGGCCACCCACACACTTGGGATCTGGGACACGAGACTTCGATTC	1179
Qy	341	LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetCylHis	360
Db	1180	AAGATGTGTCAAAAGGTACAAATGACAACTCTTGACAGCCATCAAGAGATGGGACAC	1239
Qy	361	IleGlnTyrAspMetAlaTyrAlaGlnProPheLeuLeuArgAsnGlyValAsnGlu	380
Db	1240	ATCCAAATATGACATGGCATATGCCAGGCAACCTTCTCGTAAAGAAACGGACCAATATA	1299
Qy	381	GlyPheHisGluAlaValGlyLutIleMetSerLeuSerAlaAlaThrProLysHisLeu	400
Db	1300	GGGTTTCATGAAGCTGTGGAGAAATCATGTCACTTCTTGACGCTACCCCAAGCATGTG	1359
Qy	401	LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnIle	420

LOCUS	AX047760	2415 bp	DNA	linear	PAT 15-DEC-2000
DEFINITION	Sequence 3 from Patent W00070032.				
ACCESSION	AX047760				
VERSION	AX047760.1	GI:11876767			
KEYWORDS					
SOURCE	synthetic construct.				
ORGANISM	synthetic construct.				
REFERENCE	artificial sequences.				
	1 (bases 1 to 2415)				
Db	1366	AAATGCATTTGGCTTCTGGCATCCGATTTTCAAGAAGATAGCGAAACAGAGATTAACCTC	1411		
QY	421	LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLys	440		
Db	1420	CTACTGAAACACAGCATTTGACAAATTTGTGGAGACATACCGTTACTTACTGTTAGAGAAG	1475		
QY	441	TrpArgTrpMetValPheLysGlyGlnIleProLysAspGlnTrpMetLysLysTrpTrp	460		
Db	1480	TGGAGGTGATGTCTTTCGGGGTGAATAATCCCAAGAGACATGGATGAAAAAGGGTGG	1538		
QY	461	GluMetLysArgGlnIleValGlyValGlnProValProHisAspGlnThrTyrCys	480		
Db	1540	GAGATGAAGCGGAGATGCTGTGGTGGGAGCGCTGCGTCGTGATGAACAACTACTGT	1599		
QY	481	AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg	500		
Db	1600	GACCGCTGCATCTCTTCCATGTTTCTAATGATTTCAATTCATTCGATATATACCAAGG	1655		
QY	501	ThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaIleHisGlnGlyPro	520		
Db	1660	ACCAATTTCACATTCACAGTTTCAAAACACTCTTTTCTCAAGCAGCTAACTATTAATGCTCT	1719		
QY	521	LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyGlnLysLeuPheAsnMetLeu	540		
Db	1720	CTGCACCAATGTGACATCTCAAAATTCACCTGAAGCTGGGAGAAAGTTCTCAAGATGCTG	1779		
QY	541	ArgLeuGlyLysSerGlnProThrPheLeuAlaLeuGlnGlnAsnValGlyValAlaAsn	560		
Db	1780	AGCTCTTGAANTCGACGACCCTCGACCCCAAGCCTTGGAATAATGGTGAGACCAAGAT	1839		
QY	561	MetAsnValArgProLeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuLysAspGln	580		
Db	1840	ATGGATGTGAAACCACTCTCAATTAATTCACCAACCGTTGTGACTGCTGAAAGAGCAG	1899		
QY	581	AsnLysAsnSerPheValGlyTyrSerThrAspTrpSerProTyrAlaAspGlnSerIle	600		
Db	1900	AACCAAAATTTCTTTGTGGGTGGACACTGAATGAGACCCCATATGCGCAACCAAGCATT	1959		
QY	601	LysValArgIleSerLeuLysSerAlaLeuGlnLysPheLysAlaTyrGluTrpAsnAspAsn	620		
Db	1960	AAAGTGAGATTAACGCTTAATAATCAGCTCTTGAGAGCTAATGCATATGCAATGACCAACAC	2019		
QY	621	GluMetTrpLeuPheAsnArgSerSerValAlaIleAlaMetArgGlnTyrPheLeuLysVal	640		
Db	2020	GAAATGTCTCCGTGTCGATCATGTGGTGCATATCCCATGACGAAGATATTTCTCAATATATC	2079		
QY	641	LysAsnGluMetIleLeuPheGlyGlnGlnLysPheValArgValAlaAsnLeuLysProArg	660		
Db	2080	AAAAACCGACAGATTCCTTTCTGAGAGGAATATGTACAGATGATGATTTAAACCAAGA	2139		
QY	661	IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr	680		
Db	2140	GTGTCCTCTTACTCTTCTTGTACCTCAACCCCAAAATGTGCTGATGTGCATCTCAAGAGT	2199		
QY	681	GluValGlnLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn	700		
Db	2200	GAACTTGAAGATGGCATTACAGATATCTCGGGGCCCATCAATGATGATCTTTGGCCTGAAT	2259		
QY	701	AspAsnSerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro	720		
Db	2260	GATTAACAGCCCTGGAGTTTCTGGGATTCACCAACCAACACTTGAGCCACCTTACCAAGCTCT	2319		

AUTHORS Piddington, C.S., Petrie, C.R., Shoemaker, K.E. and Bishop, P.D.
 TITLE Zace2: a human metalloenzyme
 JOURNAL Patent: WO 00/0032-A 3 23-NOV-2000;
 ZymoGenetics, Inc. (US)

FEATURES
 source
 1. .2415
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="This degenerate sequence encodes the amino acid
 sequence of SEQ ID NO:2."
 1. .2415

misc_feature

BASE COUNT 494 a 218 c 398 g 335 t 970 others
 ORIGIN

Alignment Scores:

Pred. No.: 1.3e-265 Length: 2415
 Score: 3175.00 Matches: 578
 Percent Similarity: 80.39% Conservative: 0
 Best Local Similarity: 80.39% Mismatches: 141
 Query Match: 82.06% Indels: 0
 DB: 6 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AX047760 (1-2415)

QY 2 ThrileglugluGlnAlaIleuLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeu 21
 DB 58 ACNATHGARGARARCGNARACNTTYTNGATAARTTAAATCAATGACGCGNARGAYTIN 117
 QY 22 PheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluAsnVal 41
 DB 118 TTYTAYCARWMSNMTYNTGCMWSMTGCAATTAATTAACNAATATACNGARARAAATG 177
 QY 42 GluAsnMetAsnAlaGluAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeu 61
 DB 178 CARAAATGAAATATGCGNGGNGAYAAATGCMSCNTTYTNAARARARCMASNACTYN 237
 QY 62 AlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAla 81
 DB 238 GCNCARATGTAACCTNYNCARCAATGACAAATATNACNTNARNTNARNTNARNTNAR 297
 QY 82 LeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIle 101
 DB 298 YTNACARARAYGGMSNMTNGTNTYTNWSNGARATTAARASNAARMGNTTNAATACNATH 357
 QY 102 LeuAsnThrMetSerThrIleTyrSerThrGlyValCysAsnProAspAsnProGln 121
 DB 358 YTNAAATACNATGMSNACNATHTAATWASNAACNGNARGTNTGYAACCGNAGAACAR 417
 QY 122 GluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsn 141
 DB 418 GARGGYTNTYNTYNGARCCNGGNYTNAAGARATTAATGCGNAAWMSNTYNGATTAAY 477
 QY 142 GluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuAspProLeu 161
 DB 478 GARGNNTTNTGGCGNTGGARGMSVTGGMGMSNARKTNGSNAARARARNTNMGCCTYN 537
 QY 162 TyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyr 181
 DB 538 TAYGARGARATYGTNGTNTYTAARAAGARATGCGNMGNCNAAATYATYAGARGATAY 597
 QY 182 GlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArg 201
 DB 598 GGNAGTATYATGCMGNGGATTAAGARGTNAATYAGNCTNGATYAGTATTAWSMGN 657
 QY 202 GlyLeuLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHis 221
 DB 658 GGNCAATYNTAHGARGATGARGACATACNTTARGARARATHAARCCNTYNTATAGACAY 717
 QY 222 LeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIle 241
 DB 718 YTNCAATGNTATGTMGNCNAAATYATTAATGAATYATCCNMTNTATYATHTMSNCCNATH 777

QY 242 GlyCysLeuProAlaHisLeuGluLysAspMetTrpGlyArgPheTrpThrAsnLeuTyr 261
 DB 778 GGNITTYTNCNCGNCCATYTTNTNGNCAATATGCGGCMGNTTYTGGACNAATYTTAT 837
 QY 262 SerLeuThrValProPheGlyGlnLysProAsnIleAspValIleAspAlaMetValAsp 281
 DB 838 WSNYTNACNCGNCCATYTTNGNCAARARCNAAATYATGATGTCNMGNGATGTCAT 897
 QY 282 GlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGly 301
 DB 898 CARGCNTGGAGCGCACMGATNTTAAARARCGNARARATYTTTGTGMSNGTCGN 957
 QY 302 LeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnVal 321
 DB 958 YTNCCNAAATATGACNCAARGNTTGGARAAATYATGATGTCNMGNGATGTCAT 1017
 QY 322 GlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeu 341
 DB 1018 CARAARGCNGTNTGCAATCCNACNCGTGGATYTNMGNAARGGATTYTGMATHTYN 1077
 QY 342 MetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHisIle 361
 DB 1078 ATGTGACNAARAGTNAACATGATGATYTTTNCNCGNCAATYATGARGATGCGNCAAT 1137
 QY 362 GlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGly 381
 DB 1138 CARTAYGATATGCTTAATGCGNCCARCCNTTYTNTYTNMGNAATYAGCGNCAAGARG 1197
 QY 382 PheHisGluAlaValGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLys 401
 DB 1198 TTYCAATGAGCNGTNGGARGATTAATGMSNTTMSNCGNCCNACNCAATYTTNAR 1257
 QY 402 SerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 421
 DB 1258 WGNATHGNGYTYTWTMSNCCNCAATYTCARARGAATYAGARACGARGATTAATYTYTN 1317
 QY 422 LeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrp 441
 DB 1318 YTNAAACARGCNYTNACNTHGTNGNACNTYNTCCNTTYACATATATGTNGARARATGG 1377
 QY 442 ArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpGlu 461
 DB 1378 MNTGATGCTGTTTAARGNARGARATHTCCNARGAATGATGATTAARAATGGTGGAR 1437
 QY 462 MetLysArgGluIleValGlyValGluProValProHisAspGluThrTyrCysAsp 481
 DB 1438 ATCAARMGARGATHTGTNGTNGTNGARCCNCGNCCAGAGARGARACATATYTGAY 1497
 QY 482 ProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThr 501
 DB 1498 CCNCGMSNTYTTTCATGTMNSNAATYATYATWSTYATHTMGATATTAACNMGNCN 1557
 QY 502 LeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluIleProLeu 521
 DB 1558 YTNATYCATTTTCATYTCARARGCNYNTYTCARGCGNCAACAAAGARGCGNCTYN 1617
 QY 522 HisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArg 541
 DB 1618 CAATAATGATATHTMSNAATYATWSTNACNARGCGNCCAGAAATYTTTAATATGTTNMG 1677
 QY 542 LeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValAlaGlyAlaLysAsnMet 561
 DB 1678 YTNMGNAARMSNARGCNGTNGACNTYNTNGATNGATNGATNGATNGATNGATNGATNG 1737
 QY 562 AsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsn 581
 DB 1738 AATGTMGNCNTYNTTNAATYATYATGARGCNYTNTYACNTGATYTNAAARAYCARAY 1797
 QY 582 LysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerLysLys 601
 DB 1798 AARAAATWSTYTTGTNGTNGTNGMSNACNGATYTGMSNCCATATGCGATCAARASNTTHAR 1857
 QY 602 ValArgIleSerLeuLysSerAlaLeuGluLysAspLysAlaTyrGluTrpAsnAspAsnGlu 621

```
Db      1858 GTNMGNAITHMSNTTNAARMNSGNCYTTNGNGATATAARCCNTATGARTGGAAYGAAYGAR 1917
QY      622 MetTyrLeupheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLys 641
Db      1918 ATGTAAYTNTTYMGMSMSNGTNGCNTATGCAATGMCNCARTATTTTNTNAAAGTNAAR 1977
QY      642 AsnGlnMetIleLeupheGlyGluAspValArgValAlaAsnLeuLysProArgIle 661
Db      1978 AATCAATGATHTTNTTYGGNGARGAGATGTMGNGTNGCNAAYTTNAAARCCMGNATH 2037
QY      662 SerPheAsnPhenPheValThrAlaProLysAsnValSerAspIleIleProArgThrGlu 681
Db      2038 MSNTTYAATTTTTCGNAACNGCCNAARAATGTMNSGATATHTAHCNMGMNACNGAR 2097
QY      682 ValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAsp 701
Db      2098 GTNGARARARCCNATHMGNAATGMSMNGMSMGNATHAAGAYGCNTTYMGNTTNAAYGAY 2157
QY      702 AsnSerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
Db      2158 AATMSNTTNGARTTTTNTGNGNATHCARCCNACNTNGGCCNNAAYCARCCNCCN 2214
```

Search completed: May 26, 2003, 19:37:57
Job time : 5047 secs

10

11

12

13

14

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 17:57:49 ; Search time 424 Seconds
(without alignments)
3824.148 Million cell updates/sec

Title: US-09-978-385-2_COPY_19_738
Sequence: 1 STIEQATFLDKFHEAD.....DNSLEFLGIQPTLPNNQP 720

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239.seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV-xlh
-Q-cg2.1/USPTO/US0997838/runat_21052003_154419_19824/app_query.fasta_1.903
-DB-N.Geneseq.101002 -OPMT-fastap -SUFFIX-rng -MINMATCH=0.1 -LOOPEXT=0
-LIST=45 -DOCALIGN=200 -THR_SCORE=100 -THR_MIN=0 -ALIGN=20
-MODE-LOCAL -OUTPM-PLC -NORMEXT -HEAPSIZE=500 -MINLEN=2000000000
-USER=050978385_CGN_1_1_263_runaat_21052003_154419_19824 -NCPV=6 -ICPV=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N.Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	2418	AAZ59465	Human MPROT15 codl
2	3869	100.0	3334	AAZ59465	Human Zace2 protel
3	3869	100.0	3396	AAA12764	CDNA encoding a hu
4	3869	100.0	3396	AAAD2758	Human angiotensin
5	3869	100.0	3396	ABK87622	CDNA encoding huma
6	3869	100.0	3396	AAAD32586	Human ACE-2 full-l
7	3833	99.1	2920	AAAD4880	Human CDNA encodin
8	3785	97.8	2911	AAAD4890	Human CDNA encodin
9	3720	96.1	3732	AAAD1279	Human MPROT15 codl
10	3366.5	87.0	2262	AAZ59466	Human CDNA encodin
11	3251	84.0	2638	AAZ59466	Human MPROT15 codl
12	3251	83.6	2638	AAZ59466	Human MPROT15 codl
13	3214	83.1	2350	AAZ59466	Human MPROT15 codl
14	3175	82.1	2415	AAZ59466	Human MPROT15 codl
15	2832	73.2	3474	AAZ59466	Human MPROT15 codl
16	2640	68.2	2415	AAZ59466	Human MPROT15 codl
17	1337	34.6	4020	AAZ59466	Human MPROT15 codl
18	1337	34.6	4020	AAZ59466	Human MPROT15 codl
19	1337	34.6	4020	AAZ59466	Human MPROT15 codl
20	1337	34.6	4020	AAZ59466	Human MPROT15 codl
21	1337	34.6	4020	AAZ59466	Human MPROT15 codl
22	1336	34.5	4020	AAZ59466	Human MPROT15 codl
23	1334	34.5	3939	AAZ59466	Human MPROT15 codl
24	1334	34.5	3939	AAZ59466	Human MPROT15 codl
25	1310	33.9	3942	AAZ59466	Human MPROT15 codl
26	1275	33.0	5005	AAZ59466	Human MPROT15 codl
27	1233.5	31.9	2568	AAZ59466	Human MPROT15 codl
28	1069	27.6	2089	AAZ59466	Human MPROT15 codl
29	1054	27.2	2074	AAZ59466	Human MPROT15 codl
30	1025.5	26.5	2450	AAZ59466	Human MPROT15 codl
31	959	24.8	5006	AAZ59466	Human MPROT15 codl
32	924.5	23.9	5632	AAZ59466	Human MPROT15 codl
33	919.5	23.8	5060	AAZ59466	Human MPROT15 codl
34	841	21.7	2379	AAZ59466	Human MPROT15 codl
35	721	18.6	2082	AAZ59466	Human MPROT15 codl
36	721	18.6	2082	AAZ59466	Human MPROT15 codl
37	721	18.6	2082	AAZ59466	Human MPROT15 codl
38	715.5	18.5	1395	AAZ59466	Human MPROT15 codl
39	687	15.4	446	AAZ59466	Human MPROT15 codl
40	597.5	15.4	42573	AAZ59466	Human MPROT15 codl
41	491	12.7	2025	AAZ59466	Human MPROT15 codl
42	458	11.8	1836	AAZ59466	Human MPROT15 codl
43	441	11.4	5116	AAZ59466	Human MPROT15 codl
44	431	11.1	4001	AAZ59466	Human MPROT15 codl
45	429	11.1	561	AAZ59466	Human MPROT15 codl

ALIGNMENTS

RESULT 1
AAZ59465
AAZ59465 standard; DNA: 2418 BP.
AAZ59465;
11-APR-2000 (first entry)
Human MPROT15 coding sequence #1.
MPROT15: treatment; hypertension; human; myocardial disease; apoplexy;
heart disease; apoplexy; heart disease; nervous denaturation; ds;
Alzheimer's disease; hormone; cytokine.
Homo sapiens.
XX Jp11318472-A.
XX

PD 24-NOV-1999.
 XX 22-JAN-1999; 99JP-0014949.
 XX 13-MAY-1998; 98GB-0010373.
 PR 18-AUG-1998; 98GB-0018009.
 XX (SMIK) SMITHKLINE BEECHAM PLC.
 PA WPI: 2000-109268/10.
 DR P-PSDB: AAY67310.
 XX
 XX MPR015 polypeptide and MPR015 polynucleotides - useful for the
 PT treatment of hypertension, myocardial diseases, apoplexy, heart
 PT diseases, nervous denaturation, Alzheimer's disease etc.
 XX
 XX Claim 7; Page 14; 22pp; Japanese.
 PS
 CC This is the coding sequence of human MPR015. The MPR015 polynucleotide
 CC and polypeptide sequences can be used for the treatment of hypertension,
 CC myocardial diseases, apoplexy, heart diseases, nervous denaturation,
 CC Alzheimer's disease and diseases related to the processing of peptide
 CC hormones and cytokines.
 CC
 XX Sequence 2418 BP; 744 A; 484 C; 555 G; 635 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 0 Length: 2418
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AAY59465 (1-2418)
 QY 1 SerThrIleGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 55 TCCACCATGAGGAAAGGACGACATTTTGGACAAAGTTTACACGAAAGCCGAGAC 114
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheTyrThrAsnIleThrGluGluAsn 40
 DB 115 CTGTCATATCAAAAGTCACCTGCTTGGATTATTAACACCAATATTTACTGAGAGAAAT 174
 QY 41 ValGlnAsnMetAsnAlaGluAspLysTyrPheLeuLysGluGlnSerThr 60
 DB 175 GTCCAAACATGATATATGCTGGGACAAATGCTGCTTTTAAAGAAACATCCACA 234
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 235 CTGGCCCAATGATGTCACATCAAGAAATTCAGAAATCTCACAGTCAAGCTTCAGCTCAG 294
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 295 GCTCTTCAGCAAAATGGGCTTCAGTCTCTCAGAAAGCAAGCAAAACGGTTGAACACA 354
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAspPro 120
 DB 355 ATTCTAATATCAATGAGACACATCTACAGTACTGGAAGTTTGTAAACCCAGAAATCCCA 414
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 415 CAAGAATGCTTATTAAGTTCAGACAGGTTTGAATGAATATGCAAAACAGTTTGAAGTAC 474
 QY 141 AsnGluArgLeuThrPheAlaTyrGlnSerTyrPheSerGluValAlaGlyLysGlnLeuArgPro 160
 DB 475 AATGAGAGGCTCTGGGCTTGGAAAGCTGGAGATCTAGGCTGGCAAGCAGCTGAGGCCA 534
 QY 161 LeuTyrGlnGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 535 TTATATGAGAGATATGCTGCTTGAATAATAGATGGCAAGCAAAATCATTTAGAGAGAC 594
 QY 181 TyrGlyAspTyrTyrPheGlyAspTyrGluValAlaAsnGlyValAspGlyTyrAspTyrSer 200

DB 595 TATGGGATTTATGGAGAGACATATGAAATGGGGTATGATGCTATGACATACAGC 654
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 DB 655 CGCGGCGAGTATGATGAAATGATGGAACATACCTTTGAAGATTAACATTAATATGAA 714
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 715 CATCTTCATGCTTATGAGGCGCAAGTTGATGATGCTTATCTCTTATATCATGCTCA 774
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheTyrThrAsnLeu 260
 DB 775 ATTGATGAGCTCCGCTCATTTGCTGGTGTATGATGGGGAGATTTTGGCAAAATCTG 834
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 835 TACTCTTGAACAGTCCCTTGGACAGAAACCAACATAGATGTTACTGATGCAATGGTG 894
 QY 281 AspGlnAlaTyrAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 895 GACGAGGCTGGGATGCGACAGAGATTTCAAGAGGCGGAGAAATTTCTTGTATCTGTT 954
 QY 301 GlyLeuProAsnMetThrGlnIlyPheTyrGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 955 GCTCTCTTAATATGACTCAAGAGATCTGGAAATTCATGCTTAACGAGCCAGCAAGAAAT 1014
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIle 340
 DB 1015 GTTCCAAAGAGAGTGGCATCCACAGCTTGGAGCTGGGAGAGGAGCTTCAGAGATC 1074
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHis 360
 DB 1075 CTTATGTCACAAAGTGCATATGAGCAGATTCCTGAGACAGCTCATCATGATGGGGAT 1134
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyValAsnGlu 380
 DB 1135 ATCCAGTATGATATGCAATATGCTGCACAACTTTTCTGCTAAGAAATGAGAGCTAATGAA 1194
 QY 381 GlyPheHisGluAlaValAlaGlyIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1195 GGAATTCATGAAAGCTGTTGGGAATCATGTCATCTTGGAGCCACACCTTAAGCATTTA 1254
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnIle 420
 DB 1255 AATTCATGATGGCTTCTGTCACCCGATTTTCAAGAAACAAATGAAACGAAATTAACCTTC 1314
 QY 421 LeuLeuLysGlnAlaLeuThrIleValAlaGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1315 CTGCTCAAAACAGACATCGATGTTGGAGCTTGCACATTTACTTACATGTTAGAGAG 1374
 QY 441 TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTyrTrp 460
 DB 1375 TGGAGGTGATGCTCTTTTAAAGGGAATTTCCCAAGACCATGATGAAAGAGGTGG 1434
 QY 461 GluMetLysArgGluIleValAlaGlyValAlaGluProValProHisAspGluThrTyrCys 480
 DB 1435 GAGATGAGAGAGAGATAGTGGGGTGGTGAACCTGTGCCCATGATGAAGAAATACATGCT 1494
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArg 500
 DB 1495 GACCCGCGAGTCCTGTCATGTTCTAATGATTCCTCATTCATTCGATATTAACCAAG 1554
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluIlePro 520
 DB 1555 ACCCTTACCAATTCAGTTTCAAGAGACCTTGTCAAGCAGCTTAACATGAGAGCCCT 1614
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGluGlnLysLeuPheAsnMetLeu 540
 DB 1615 CTGCACAAATGTGACATCTCAAACTCTACAGAGGCTGGACAGAAATCTGTCAATATGCTG 1674
 QY 541 ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValAlaLysAsn 560


```

Db 1675 AGGCTTGAAATTCAGAACCCGACCTAGATTGGAAATGTTGTAGACCAAGAAC 1734
QY 561 MetasnaValArgProLeuLeuasnTyrrPheGluProLeuPheThrTrpLeuLysAspGln 580
Db 1735 ATGAATGTAAAGCCACCTGCTCAACTTGTGAGCCCTTATTTACCTGGCTGAAAGCCAG 1794
QY 581 AsnLysAsnSerPheValGlyTrpSerThrAspTyrPseProTyrAlaAspGlnSerIle 600
Db 1795 AACCAAAATCTTTGTTGGATGGAGATACCGACTGGAGCCATATGATGCAATGGAAGCAATC 1854
QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn 620
Db 1855 AAGTAGAGATAGCTTAATATCAGCTCTTGAGAGATTAAGCATATGATGGAAGCAATC 1914
QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
Db 1915 GAAATGTACCTGCTTCGATCATCTGTGCATATGATATGATGAGCAGACTTTTAAAGTA 1974
QY 641 LysAsnGluMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
Db 1975 AAAAATCAGATATCTTTTGGGAGAGAGATGCGCAGTGGCTAATTTGAAACCAACA 2034
QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
Db 2035 ATCTCCTTAATTTCTTTGTCTACTGACCTAAATATGCTCATATCTTCTAGACT 2094
QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
Db 2095 GAAGTTGAAAGAGCCATCAGATGATGCCGAGCCGCTATCAATGATGCTTCCGCTGCAAT 2154
QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
Db 2155 GACAAACAGCTAGAGTTCTTGGGAGATACAGCCAAACATCTGCTTAACAGCCCCCT 2214
RESULT 2
AAC84366
ID AAC84366 standard; cDNA; 3334 BP.
XX
AC AAC84366;
XX
DT 19-MAR-2001 (first entry)
XX
DE Human Zace2 protein encoding cDNA.
XX
KW Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
KW zinc metalloproteinase; blood pressure; zinc protease; hypertension;
KW ventricular systolic dysfunction; renal impairment; heart failure;
KW scleroderma renal crisis; atherosclerosis; antiinflammatory; human;
KW antirheumatic; bradykinin inactivator; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 35..2452
FT /tag="a
FT /product="Zace2"
XX
XX W0200070032-A1.
XX
XX PD 23-NOV-2000.
XX
XX PF 03-MAY-2000; 2000MO-US11932.
XX
XX PR 13-MAY-1999; 99US-0311482.
XX
XX PR 27-AUG-1999; 99US-0384706.
XX
XX PA (ZYMO) ZYMOGENETICS INC.
XX
XX PI Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
XX
XX WPI; 2001-025018/03.
XX
XX DR P-PSDB; AAB48095.
XX

```

```

PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
PT associated with inflammation such as arthritis and enterocolitis -
XX
XX Example 1; page 95-100; 125pp; English.
XX
CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-
CC converting enzyme is a zinc metalloproteinase that plays roles in blood
CC pressure regulation and fertility. Zace2 can be expressed by standard
CC recombinant methodology. Zace2 polypeptides are useful for treating an
CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
CC diseases associated with inflammation like arthritis and enterocolitis,
CC as targets for identifying new angiotensin-converting enzyme (ACE)
CC inhibitors, and as a basis for rational drug design for inhibitory
CC molecules. The nucleic acids can be used to detect the expression of a
CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and
CC for detecting and localizing Zace2 gene expression in tissue samples,
CC to determine whether a subject's chromosomes contain a mutation in the
CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.
CC Inhibitors of ACE are used for treating hypertension of various
CC conditions, including left ventricular systolic dysfunction, progressive
CC renal impairment, scleroderma renal crisis, congestive heart failure due
CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
CC used to treat infertility while Zace2 antagonists are used for inducing
CC infertility. The present sequence represents a cDNA encoding the human
CC Zace2 protein.
XX
SQ Sequence 3334 BP; 1011 A; 640 C; 754 G; 929 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0 Length: 3334
Score: 3869.00 Matches: 720
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0
XX
US-09-978-385-2_COPY_19_738 (1-720) x AAC84366 (1-3334)
QY 1 SerPrrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
Db 89 TTCACCATTTGAGAGAACAGGCCAAGACATTTTGGACAAATTTAAACCAAGAGCCGAAGAC 148
QY 21 LeuPheTyrGlnSerSerIleuAlaSerTyrAsnTyrAsnThrAsnIleThrGluGluAsn 40
Db 149 CTGTTCTATCAAAAGTCTACTGCTCTTGGAAATTAATACCAATATTACTGAAGAGAAAT 208
QY 41 ValGlnAsnMetAsnAlaGlyAspLysTyrSerAlaPheLeuLysGluGlnSerThr 60
Db 209 GTCCAAAACATGAAATATGCTGTGGGAGACAAATGCTGCTCTTTTAAAGAAAGATGCCACA 268
QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuPheValLysLeuGlnLeuGln 80
Db 269 CTTGCCCAATATGTTCCACTACAGAAATTCAGATCTCACAGTCAGCTTCACTGCTCAG 328
QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
Db 329 GCTCTTACAGCAAAATGGGTCTTCAAGTGTCTCAGAACAAAGACAAACGGTTGAACACA 388
QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnPro 120
Db 389 ATTCTTAATACAAATGAGACCATCATCTACACTGTGAAAGATTTGTATACCCAGATTAATCCA 448
QY 121 GlnGluCysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyr 140
Db 449 CAAAGATCTTATTACTTGAACCAAGGTTTGAATATAATGGCAACACAGTTTGACTAC 508
QY 141 AsnGluArgLeuTyrPalatPgluSerTyrArgSerGluValGlyLysGlnLeuArgPro 160
Db 509 AATGAGAGAGCTCTGCGGCTTGGGAAAGCTGAGATCTGAGGTGCGCAAGACGATGAGGCCA 568
QY 161 LeuTyrGluGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAsp 180

```

Dd	569	TTATTTGGAAGAGTATGCTGCTCTTGAATAATGAGATGGCAAGCAAGCAATCATTTATGAGGAC	628
Qy	181	TyrGlyAspTyrTrrArgGlyAsprrcrluValAsnGlyValAspGlyTyrAsprrrser	200
Dd	629	TATGGGGATTTATGGAGAGAGACTATGAAATGTAATGGGGTGAAGTGGCTATGACTACACGC	688
Qy	201	ArgGlyGlnLeuIleGluAspValGlnIstThrPheGlnGlnIleLeuPheLeuTyrGlu	220
Dd	689	CGGGCCAGTGTGATGTAAGATGTGGACATACCTTTGAAGAAGATTAAACCATTTATATGAA	748
Qy	221	HisLeuHisAlaIstYrValArgAlaLysLeuMetAsnAlaIstYrProSerTyrIleSerPro	240
Dd	749	CATCTTCCATGGCTTATGTAGTGGCCAAAGTGAATGGCTTATCCCTTCTATATACAGTCCA	808
Qy	241	IleGlyCysLeuProAlaHisLeuLeuGlnLysAspMetTrrGlyArgPheTrrThrAsnLeu	260
Dd	809	ATTGGATGCCCTCCCGTCATTTGGTGTGTATGTATGTGGGAGATTTTGGACAATGTG	868
Qy	261	TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal	280
Dd	869	TACTTTTGACAGTTCCTCTTGGACAGAAACCAACATATGTTACTAGATGCATGTGTG	928
Qy	281	AspGlnAlaTrrPaspAlaGlnArgIlePheLysGlnIstGlnLysPhePheValSerVal	300
Dd	929	GACCAAGCCCTGGGATGCACAGAGATATTCAAGAGGCCGAGAAATGTTGTATCTGTT	988
Qy	301	GlyLeuProAsnMetThrGlnGlyPheTrrGlnAsnSerMetLeuThrAspProGlyAsn	320
Dd	989	GGTCTTCCATATATGATCACAAGGATTTCTGGGAAATTCATAGTCCTAACGACCAGCAAT	1048
Qy	321	ValGlnLysAlaValCysHisProThrAlaTrrPaspLeuGlyLysGlyAspPheArgIle	340
Dd	1049	GTTCCGAAGAAGTGTGCCATCCCACTTGGGACCTGGGAGAAAGGGAGACTTCAGATTC	1108
Qy	341	LeuMetCysTrrLysValThrMetAspPheLeuTrrAlaHisHisGlnIleMetGlyHis	360
Dd	1109	CTTATGTGCACAAAGGTGACAAATGAGACACTTCCGTACAGCTCATCATGAGATGGCGAT	1168
Qy	361	IleGlnTyrAspMetAlaTrrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGln	380
Dd	1169	ATCCAGTrrrGATATGCGATATGCTGCCAACACTTTCCTGCTTAAAGAAATGGAGCTAATGAA	1228
Qy	381	GlyPheHisGlnAlaValGlnIleLeuMetSerLeuSerAlaAlaThrProLysHisLeu	400
Dd	1229	GGATTCCTAGAAAGCTGTGGGGAATCATGTGCATCTTGTGCAGCCACACCTTAGACATTTA	1288
Qy	401	LysSerIleGlyLeuLeuSerProAspPheGlnGlnLysAsnGlnIstGlnIleAsnPhe	420
Dd	1289	AAATTCATTTGGCTCTGTCTGCACCCCATTTTCAAGAGACAATGAAACAGAAATTAACCTTC	1348
Qy	421	LeuLeuLysGlnAlaLeuThrIleValGlnIstThrLeuProPheThrTyrMetLeuGlnLys	440
Dd	1349	CTGCTCCAACAAGCACTCACATGTTGTGGACTGTGCCATTTACTTACATGTrrrTAGAGAG	1408
Qy	441	TrrArgTrrMetValPheLysGlnIstLeuProLysAspGlnTrrMetCysLysTrrPrr	460
Dd	1409	TGGAGCTGGATGGTCTTTAAAGGGGAATTCOCAAGAAGCAAGTGGATGAATAAAGTGGTG	1468
Qy	461	GlnMetLysArgGlnIstIleValGlnIstValAlaGlnProValProHisAspGlnThrTyrCys	480
Dd	1469	GAGATGAAGACGAGAGATAGTGGGTGTGGAACTGTGCCCATGATGAAGAAACATACGT	1528
Qy	481	AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrrThrArg	500
Dd	1529	GACCCCGCATCTCTGTTCCATAGTrrrCTATATGATTACTATTCATTCGATTTACACAAGG	1588
Qy	501	ThrLeuTyrGlnPheGlnPheGlnGlnIstLeuLysCysGlnAlaAlaLysHisGlnIstPro	520
Dd	1589	AACCTTTTACCAATTCACATTCACAAGAAGCACTTTGTCAAGCAGTAAACATGAAGGCCCT	1648
Qy	521	LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlnIstGlnLysLeuPheAsnMetLeu	540

Db	1649	CTGGACAAATGTGCATCTCAAACTCTACAGAAACCTGGACAGAAACTGTTCAATATCTGTG	1708
QY	541	ArgLeuGlyLysSerGlnProTyrThrLeuAlaLeuGluAsnValValGlyAlaLysAsn	560
Db	1709	AGGCTGGAAATATCGAACCTCGACCTGACCATTTGGAAATGTGTGAGCAAGAAC	1768
QY	561	MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTriPheLeuLysAspGln	580
Db	1769	ATGAATGTAAAGGCCACTGCTCAACTACTTTGAGCGCCTATTATTAACCTGGCTGAAGACAG	1828
QY	581	AsnLysAsnSerPheValGlyTyrSerThrAspTrpSerProCysAlaTyrGluTrpAsnAspAsn	600
Db	1829	AACAGAAATCTTTTGTGTGGATGGATGACGATCCAGCTGGAGTCCATATGCAACCAAGCATC	1888
QY	601	LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn	620
Db	1889	AAAGTAGAGATTAAGGCTAAATACAGCTCTTGAGATGAATGAATGAGAACAGACAT	1948
QY	621	GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal	640
Db	1949	GAAATGTATCCCTGTTCGATCATCTGTTCGCAATCTATAGAGCGAGTACTTTTAAAAAGTA	2008
QY	641	LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg	660
Db	2009	AAAAATCGATGATTTCTTTTGGGAGAGAGATGTCCAGTGGCTAAATTTAAACCAGCA	2068
QY	661	IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr	680
Db	2069	ATCTCCCTTAATTTCTTGTGTGCTGACCTGACCTAAATATGCTGTGATATCATCTTAGA	2128
QY	681	GluValGlyLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn	700
Db	2129	GAGTGTGAAAAAGCCATCAGGATGTCCCGGAGCCGTATCAATGATGCTTCGCTGAAT	2188
QY	701	AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro	720
Db	2189	GACAACAGCCTTAGAGTTCTTGCGGATACAGCCAAACACTTGAGCCTCTCCAAACACCCCT	2248
RESULT 3			
AA12764			
ID	AA12764	standard; cDNA: 3396 BP.	
XX	AA12764;		
AC			
XX	25-JUL-2000	(first entry)	
DT			
DE	cDNA encoding a human angiotensin converting enzyme-2 (ACE-2).		
XX			
XX	Human; angiotensin converting enzyme-2; ACE-2; angiotensin I; Ang. (1-9);		
KW	blood pressure; hypertension; congestive heart failure; atherosclerosis;		
KW	chronic heart failure; acute heart failure; myocardial infarction;		
KW	renal failure; SS.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	82..2499	
FT		/*tag= a	
FT		/product= "angiotensin converting enzyme-2"	
FT	sig_peptide	82..135	
FT		/*tag= b	
PN	WO200018899-A2.		
PD	06-APR-2000.		
XX			
PF	29-SEP-1999;	99WO-US22976.	
XX			
PR	30-SEP-1998;	98US-0163648.	
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
PT	Acton LS, Robison KE, Hsieh FY;		

XX WPI: 2000-293140/25.
 DR P-PSDB: AAY84562.
 XX
 PT Isolated nucleic acid encoding angiotensin converting enzyme-2 (ACE-2)
 PT polypeptide useful for detecting an ACE-2 therapeutic for treating
 PT hypertension, congestive heart failure, myocardial infarction,
 PT atherosclerosis and renal failure -
 XX
 PS Claim 1; Fig 1; 138bp; English.
 CC The present sequence encodes a human angiotensin converting enzyme-2
 CC (ACE-2). ACE-2 is expressed predominantly in kidneys and testis. The
 CC sequence of the full length ACE-2 cDNA was determined from a clone
 CC obtained from a cDNA library prepared from mRNA of a human heart of
 CC a subject who had congestive heart failure. ACE-2 has significant
 CC sequence homologies with ACE enzymes, and has also been shown to
 CC hydrolyse angiotensin I into Ang (1-9). The ACE-2 therapeutics are
 CC used to treat blood pressure related diseases and conditions, such as
 CC hypertension, congestive heart failure, chronic heart failure, acute
 CC heart failure, myocardial infarction, atherosclerosis and renal
 CC failure.
 CC
 XX
 SQ Sequence 3396 BP; 1034 A; 660 C; 771 G; 931 T; 0 other:
 Alignment Scores:
 Pred. No.: 0 Length: 3396
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AAA12764 (1-3396)
 QY 1 SerThrIleGlnGlnGlnAlaGlnThrPheLeuAspLysPheAsnHisGlnIleAsp 20
 DB 136 TCACACATTGAGAACGCGCAAGACATTTTGGACAGATTAAACACGAGAGCGGAAGAC 195
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheAsnTyrAsnThrIleThrGlnIleAsn 40
 DB 196 CTGTTCATCAAAAGTTCACCTGCTCTCTGGAATTATACACCAATTTACTGAGAGGANT 255
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerAlaPheLeuLysGlnIleSerThr 60
 DB 256 GTCCAAACATGATATATGCTGGGACAAATGCTGCTTTAAAGCAACGTCACCA 315
 QY 61 LeuAlaGlnMetTyrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnIleGln 80
 DB 316 CTGGCCAAATGATCCACTACAAAGAAATTCAGAAATCTCACAGTCAAGCTTCAGTGCAG 375
 QY 81 AlaLeuGlnIleAsnGlnSerSerValLeuSerGlnAspLysSerLysArgLeuAsnThr 100
 DB 376 GCTCTTCAGCAAAATGGCTCTTCAGTCTCTCAGAGACAAAGCAAGAGCTTGAACACA 435
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyValLysAsnProAspAsnPro 120
 DB 436 ATCTTAATATCAATGAGACCACTACATCACTAGTGAAGAAAGTTGTAAACCCAGATTAATCCA 495
 QY 121 GlnGlnLysLeuLeuLeuGlnProGlyLeuAsnGlnIleLeuAlaAsnSerLeuAspTyr 140
 DB 496 CAGAAATGCTTAATCTTACTGAAACAGGTTGAAATGAATTAATGCAACACTTTAGACTAC 555
 QY 141 AsnGlnLysLeuTyrPheAlaTyrGlnIleSerTyrPheSerGlnValGlyLysGlnLeuAspPro 160
 DB 556 AATGAGAGGCTGTGGCTTGGGAAGCTGAGATCTAGAGTGGCAAGACAGCTGAGGCCA 615
 QY 161 LeuTyrGlnGlnLysValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAsp 180
 DB 616 TTTATATCAAGATATGCTGCTTGAAGAAATGAGATGCAAGAGCAATCTATTATGAGGAC 675
 QY 181 TyrGlnAspTyrTyrPheGlnAspTyrGlnValAsnGlnValAspGlyTyrAspTyrSer 200

DB 676 TATGGGCTTTTGGAGAGGAGACTATCACTAAATGGGCTAGTGGCTATGACTACAGC 735
 QY 201 ArgGlyGlnLeuIleGlnLysValGlnIleSerThrPheGlnIleLysProLeuTyrGln 220
 DB 736 CGCGCCAGGTTGATGAAGATGTGAACATACCTTGAAGATTAACCATTTATGAA 795
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 796 CATCTTCATGCTTATGATGAGGCAAAAGTTGATGATGCTTATGCTTATGATGCA 855
 QY 241 IleGlyLysLeuProAlaHisLeuLeuGlnLysAspMetTyrPheTyrPheTyrThrAsnLeu 260
 DB 856 ATTGATATGCTTCCCTGCTGCTATTTGCTTGTATATGCTGCTATTTGGACAAATCTG 915
 QY 261 TyrSerLeuThrValProPheGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 916 TACCTTTGACAGTTCCTTTGGACAGAAACCAATAGATGTTACTGATGATGATGCTG 975
 QY 281 AspGlnAlaTyrPheAspAlaGlnArgIlePheLysGlnIleGlnLysPhePheValSerVal 300
 DB 976 GACACAGGCTGGGATGACAGAGAAATTCAGAGGCGGAGAGTCTTGTATCTGTT 1035
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrPheGlnAsnSerMetLeuThrAspProGlyAsn 320
 DB 1036 GGTCTTCTTAATATATGATCACTCAAGATTTCTGGGAAATTTCCATGCTTACAGGACCCAGAAAT 1095
 QY 321 ValGlnLysAlaValLysHisProThrAlaTyrPheLysGlnLysGlyAspPheArgIle 340
 DB 1096 GTTCAGAAAGCAGTCTGCCATCCACAGCTTGGAGCCTGGGAGAGCGGCACTTACAGATC 1155
 QY 341 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisGlnMetGlyHis 360
 DB 1156 CTATATGTCAGAAAGTGACAAATGACATGTCGACAGCTTCATCAATGATGAGGAT 1215
 QY 361 IleGlnTyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsnGlnValAsnGln 380
 DB 1216 ATCCAGATGATATGATGATATGCTGACACACTTTTCTGCTTAAGAAATGAGACTATGAA 1275
 QY 381 GlyPheHisGlnValValGlyIleLeuMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1276 GGATTCACAGAGGCTTGGGAAATTCATGTCATCTTTCGACACCACTTAAGCATTTA 1335
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnIleAspAsnGlnIleAsnPhe 420
 DB 1336 AATCCATGTTGCTTCTGACCCATTTTCAAGAGACAAAGAAACGAAATTAACCTTC 1395
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLys 440
 DB 1396 CTGCTCAAAACAGCACTACGATGTGTGGACTCTGCCATTTACTTACATGTAGAGAG 1455
 QY 441 TyrPheTyrMetValPheLysGlyLysIleProLysAspGlnTyrMetLysLysTyrPhe 460
 DB 1456 TGGAGGTGATGCTCTTAAAGGGAATTTCCCAAGACACAGTGGATGATTAAGAGGTGG 1515
 QY 461 GlnMetLysArgGlnIleValGlyValValGlnProValProHisAspGlnThrTyrCys 480
 DB 1516 GAGATGAAACGAGAGATAGTGGGTGGGAACTGTGCCATGATGAAGAAACATACAGT 1575
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArg 500
 DB 1576 GACCCCGCATCTCTGTTCATGCTTCTATATGATTAATCTCATTCATTCGATATACCAAG 1635
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGlnIlePro 520
 DB 1636 ACCCTTTACCAATTCAGTTTCAAGAGCACTTTTCAAGCGACGCTTAACATGAAGCCCT 1695
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyGlnLysLeuPheAsnMetLeu 540
 DB 1696 CTGCACAAATGTCATCTCAAACTCTACAGAAAGCTGGACAAACCTGTTCAATATGCTG 1755
 QY 541 ArgLeuGlnLysSerIleProTyrPheLeuAlaLeuGlnLysValValGlyAlaLysAsn 560
 DB 1756 AGGCTTGGAATTCAGAACCCCTGAGCCCTTGAAGTGTGTTGAGAGCAAAAGAAC 1815

QY 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrLeuLysAspGln 580
 DB 1816 ATGAATGTAAGCCACGCTCACTACTTGGAGCCCTATTACTGCGCAAGACAG 1875
 QY 581 AsnLysAsnSerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIle 600
 DB 1876 AACAGAAATCTCTTTGGGATGGAGTACGACACGAGGATGCTCATATGACACCAAGCATC 1935
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPAsnAspAsn 620
 DB 1936 AAATGAGAGTAACCTTAATATCAGCTCTTGGAGATTAACATATGATGAAAGACAAAT 1995
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
 DB 1996 GAAATGTACCTGTTCCGATCATCTGTGCATATGCTATGAGGAGGAGTCTTTTAAAGTA 2055
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGlnAspValArgValAlaAsnLeuLysProArg 660
 DB 2056 AAAAATCAGATGATCTTTTGGGAGAGAGATGTGGAGTGGCTAATTTGAAACCAAGA 2115
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 DB 2116 ATCTCTTATTTCTTCTTGCACGCGCACTAAATGTCTGATATCATCTTAGACT 2175
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 DB 2176 GAATGTAAAGGCGCATCAGAGATGTCGCCGAGCCGATCATATGATGCTTCCGTCGAAT 2235
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2236 GACACACGCTTAGATTCTTGGGATACAGCCAAACTTGGACCTCTTACACGACCCCT 2295
 RESULT 4
 AAD02758 standard; cDNA; 3396 BP.
 ID AAD02758;
 AC AAD02758;
 XX 31-MAY-2001 (first entry)
 DT 31-MAY-2001 (first entry)
 XX Human angiotensin converting enzyme-2 (ACE-2) cDNA.
 DE Human angiotensin converting enzyme-2; ACE-2; peptidyl dipeptidase A;
 KW screening; therapy; hypertension; congestive heart failure; CHF;
 KM inflammation; pain; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 82..2499
 FT /tag- a
 FT /product- "Human angiotensin converting enzyme-2
 FT (ACE-2)"
 FT /EC_number- "3.4.15.1"
 FT /note- "this region is referred as SEQ. ID. NO.3 and is
 FT specifically claimed in claim 26"
 FT sig-peptide 82..135
 FT /tag- b
 FT mat-peptide 136..2496
 FT /tag- c
 FT /product- "Human mature angiotensin converting enzyme-2
 FT (ACE-2)"
 XX US6194556-B1.
 XX 27-FEB-2001.
 XX 11-DEC-1997; 97US-0989299.
 XX 11-DEC-1997; 97US-0989299.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA

XX
 PI Acton SL, Robison KE.
 XX
 DR WPI, 2001-210604/21.
 DR P-PDB; AAY72667.
 XX
 PT Novel genes encoding angiotensin converting enzyme-2 useful as
 PT antiseize or antigenic agents for therapeutics, diagnostics and
 PT screening assays.
 PS Claim 1; Fig 1; 76pp; English.
 XX
 CC The present sequence is human angiotensin converting enzyme-2 (ACE-2)
 CC cDNA. ACE is also referred as peptidyl dipeptidase A. Nucleic acid
 CC sequence encoding ACE-2 is useful as antiseize or antigenic agents for
 CC sequence specific modulation of gene expression or in the analysis of
 CC single base-pair mutations in the gene. Nucleic acid sequence encoding
 CC ACE-2 is useful in therapeutics, diagnostics and in screening assays.
 CC ACE-2 antagonist is used to treat hypertension or congestive heart
 CC failure (CHF). ACE agonist is used to reduce the inflammation and pain
 CC resulting from an insect sting or bite, which was accompanied by an
 CC injection of bradykinin. Anti ACE-2 antibodies are used to monitor ACE-2
 CC protein levels for determining the disease or condition associated with
 CC an aberrant protein level.
 XX
 SQ Sequence 3396 BP; 1034 A; 659 C; 772 G; 931 T; 0 other;
 Alignment Scores:
 Pred. No.: 0 Length: 3396
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AAD02758 (1-3396)
 QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 136 TCCACCAATGAGAGACAGAGCCAGCAATTTTGGCAAGTTTAACCAAGACCGGAGAC 195
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrAsnTyrAsnThrAsnIleThrGluGluAsn 40
 DB 196 CTGTTCTATCAAAAGTCTCACTCTTCTTGGATATTAACCAATATATTAGAGAGAAAT 255
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerIlePheLeuLysGluGlnSerThr 60
 DB 256 GTCCAAAACATGAAATATGCTGGGACAAATGCTGCTTTTAAAGAAAGACGTCACA 315
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 316 CTGGCCCAAAATGATATCCATCAAGAAATTCAGATGTCACAGTCACAGCTTACCTCAG 375
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 376 GCTCTTCAGCAAAAGGCTCTTCAAGTGTGTCACAAAGACAGACCAACGCTTGAACACA 435
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnPro 120
 DB 436 ATCTCAATACAAAGAGAGCCATCTACAGTCTGAAAAGTTTGAACCAAGATATCCA 495
 QY 121 GlnGluLysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 496 CAAGATGCTTATTAATCTTGAACCAAGGTTTGAATGAATATGCAAAACAGCTTAAAGTAC 555
 QY 141 AsnGluArgLeuTyrPalaTyrPalaTyrSerTyrArgSerGluValGlyGlnLeuArgPro 160
 DB 556 AATGAGAGGCTCTGGGCTTGGGAAAGCTGAGATCTAGGTCGGCAAGAGCTGAGGCCA 615
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 616 TTATATGAGAGTATGTGGTCTTGAATAAATGAGATGCAAGAGCAATATTAATGAGAC 675

QY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 DB 676 TATGGGATATATGGAGAGAGACATAGTAAGTAAGGGATAGATGGCTATGACTACAC 735
 QY 201 ArgGlyGlnLeuLeuGlnAspValGlnIsthrPheGlnGlnIleLeuProLeuTyrGlu 220
 DB 736 CGGGCCAGTGTGATTAAGATGGAACATACCTTGAAGAGATTAAACCATTAATGAA 795
 QY 221 HisLeuHisAlaTyrValArgAlaLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 796 CATCTTCATGCCATATGGAGGCGCAAAAGTATGAAATGCCATCCATATACATCA 855
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeu 260
 DB 856 ATTGAGATGCCCTCCCTGCTCATTTGCTGGATATATGGGGTGGATTTGGACAAATCTG 915
 QY 261 TyrSerLeuThrValProPheGlyGlnIleProAsnIleAspValIsthrAspAlaMetVal 280
 DB 916 TACTCTTGGACATGCCCTTGGACAGAAACCAACATAGATGTACTGATGCAATGGTG 975
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLeuGlnAlaGluLeuPhePheValSerVal 300
 DB 976 GACACAGCCCTGGAGATGCACAGACATATTCAGAGAGCCGAGAGAGTCTTGTATCTGTT 1035
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 1036 GGCTCTTCCATATGACTCAAGATCTCGGAAATTTCCATCTCAACGACCCAGAGAAAT 1095
 QY 321 ValGlnIleValAlaValCysHisProThrAlaTrpAspLeuGlyLeuGlyAspPheArgIle 340
 DB 1096 GTTCAGAAAGCATCTGCCATCCACACCTTGGACCTGGGAGAGGCGCTTCAGAGATC 1155
 QY 341 LeuMetCysThrIleValThrMetAspPheLeuThrAlaHisIsthrGlnMetGlyHis 360
 DB 1156 CTATATGCGCAAAAGGAGCAATGAGAGACTTCTGACACTCATCATGATGAGGGGCAT 1215
 QY 361 IleGlnIsthrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1216 ATCCAGATATGATGGCATATGCTGCACAACTTTCTGTAAGAAATGAGCATATGAA 1275
 QY 381 GlyPheHisGlnAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLeuHisLeu 400
 DB 1276 GGATTCATGAAAGCTGTTGGGGAATCATCTCTTCTGACCCACACTTAAGCATTTA 1335
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 DB 1336 AAATCCATGTGTTCTGTCACCGCATTTCAAGAAAGCAATGAAACAGAAATAACTTC 1395
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1396 CTGCTCAAAACAGACACACAGATGTTGGACCTCTGCCATTTACTTACATGTTAGAGAAC 1455
 QY 441 TrpArgTrpMetValPheLysGlyGlnIleProLysAspGlnTrpMetCysLysTrpTyr 460
 DB 1456 TGGAGGTGGATGCTTTTAAAGGGAAATTCACCAAAACAGAGTGGATGAAAGAGTGTGG 1515
 QY 461 GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCys 480
 DB 1516 GAGATGAAGCAGAGATAGTTGGGGTGGTGAACCTGCCCCATGATGAACATATCTGT 1575
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArg 500
 DB 1576 GACCCGCAATCTGTTCCATGTTCTTAATCATTAACATTCATTCATATTTACACAGG 1635
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 DB 1636 ACCCTTACCAATTCAGATTCAGAAAGCACTTTGTACAGCAGCTAAACATGAAGGCCCT 1695
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
 DB 1696 CTCACAAATGTGACATCTCAAACTACAGAAAGCTGACAGAAACATGTTCAATATCTGT 1755
 QY 541 ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValGlyAlaLysAsn 560

DB 1756 AGGCTGGAAATATCGAACCCCTGGACCTACATTTGGAATGTTGAGGCAAAAC 1815
 QY 561 MetAsnValArgProLeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuLysAspGln 580
 DB 1816 ATGAATGTAAGCCCACTGCTCACTACTTGAAGCCCTTATTTACTGGCTAAAGACACAG 1875
 QY 581 AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIle 600
 DB 1876 AACAGAAATTTCTTTGTGGATGAGTACCGCATCGATGCCATCATATGACAGCAAAAGCATC 1935
 QY 601 LysValAlaGlyIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn 620
 DB 1936 AAAGTGAGATTAAGCTTAAACATCACTCTTGAGATTAAGCATTAATGAGACGACAT 1995
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
 DB 1996 GAAATGTACCTTCTCCGATCATCTGTGCAATATGCTATGAGGACGACTTTTAAAGTA 2055
 QY 641 LysAsnGlnMetIleLeuPheGlyGlnGluAspValArgValAlaAsnLeuLysProArg 660
 DB 2056 AAAATGACATGATTTCTTTTGGAGAGAGATGTGCGAGTGGCTAATTTGAAACCAAGA 2115
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 DB 2116 ATCTCTTTAAATTTCTTTGTCTACTGACCTAAATAATGTCTGATATCATTCAGAACT 2175
 QY 681 GluValGlnLysAlaIleArgMetSerArgIleAsnAspAlaPheArgLeuAsn 700
 DB 2176 GAAGTTGAAAGGCCATCGATCGATGCCGAGCCGATATCATGATGCTTCCGTCTGAT 2235
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2236 GACAACAGCGCTAGATTTCTGGGATACAGCCACACATTTGACCTCTAACCAGCCCT 2295
 DB 2295

RESULT 5
 ABR87623
 ID ABR87623 standard; cDNA: 3396 BP.
 AC ABR87623;
 XX
 XX
 DT 24-SEP-2002 (first entry)
 DE cDNA encoding human angiotensin converting enzyme-2 (ACE-2) protein.
 KW Human: angiotensin converting enzyme-2; ACE-2; body weight disorder;
 KW muscle mass; body fat; obesity; diabetes; atherosclerosis; weight loss;
 KW lipid metabolism; weight gain; anorexia; cachexia; bulimia; sepsis;
 KW familial partial lipodystrophy; hypercholesterolaemia; hyperlipidaemia;
 KW aberrant metabolic rate; heart failure; left ventricular hypertrophy;
 KW neurodegenerative disorder; peptide hormone; cytokine processing;
 KW myocardial infarction; cardiomyopathy; inflammatory bowel disease;
 KW systemic inflammation response syndrome; polytrauma; pain; stroke;
 KW bone destruction; rheumatoid arthritis; osteoarthritis; asthma;
 KW periodontal disease; dysmenorrhoea; premature labour; brain oedema;
 KW focal injury; diffuse axonal injury; reperfusion injury; scar formation;
 KW cerebral vasospasm; subarachnoid haemorrhage; allergic disorder;
 KW adult respiratory distress syndrome; wound healing; appetite;
 KW body mass index; gene; ss.
 XX
 OS Homo sapiens.
 FH
 FH key Location/Qualifiers
 FT CDS 82..2499
 FT /tag= a
 FT /product= "human angiotensin converting enzyme-2 (ACE-2)
 FT /tag= b
 FT /tag= d
 FT mat.peptide 82..138
 FT /tag= b
 FT 139..2496
 FT /tag= c
 XX
 PN W0200239997-A2.

XX 23-MAY-2002.
 PD 31-OCT-2001; 2001WO-US45703.
 PF 01-NOV-2000; 2000US-0704216.
 PR 29-MAY-2001; 2001US-0870382.
 PR 19-OCT-2001; 2001US-371741P.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Acton ST, Ocala FL, Gould AE, Dales NA, Guan B, Brown JA;
 PI Patane M, Kadambi VU, Solomon M, Stricker-Krongrad A;
 XI WPI: 2002-547572/58.
 DR P-PSDB: AAU99701.
 PT Treating body weight disorder and increasing muscle mass comprises
 PT administering angiotensin converting enzyme-2 modulating compound
 PS Disclosure: Page 383-387; 395pp; English.
 XX The present invention describes a new method of treating a body weight
 CC disorder, increasing muscle mass and decreasing body fat by
 CC administration of angiotensin converting enzyme (ACE)-2 modulating
 CC compound. The invention can be used for treating body weight disorders,
 CC particularly obesity of at least grade I, diabetes, atherosclerosis and
 CC a state associated with lipid metabolism. The method is used for treating
 CC rapid weight loss, rapid weight gain, anorexia, cachexia, bulimia,
 CC generalised partial lipodystrophy, familial partial lipodystrophy,
 CC hypercholesterolemia, hyperlipidaemia, an aberrant metabolic rate,
 CC congestive heart failure, chronic heart failure, left ventricular
 CC hypertrophy, acute heart failure, neurodegenerative disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease and Huntington's disease),
 CC diseases associated with peptide hormones or cytokine processing,
 CC myocardial infarction, cardiomyopathy, systemic inflammation response
 CC syndrome, sepsis, polytrauma, inflammatory bowel disease, acute and
 CC chronic pain, bone destruction in rheumatoid arthritis and osteoarthritis
 CC and periodontal disease, dysmenorrhoea, premature labour, brain oedema
 CC following focal injury, diffuse axonal injury, stroke, reperfusion
 CC injury, cerebral vasospasm after subarachnoid haemorrhage, allergic
 CC disorders including asthma, adult respiratory distress syndrome, wound
 CC healing and scar formation. The invention decreases the appetite,
 CC increases muscle mass and decreases body fat of subject having body mass
 CC index of greater than 23 (preferably 24.9)kg/m². The present nucleic
 CC acid sequence encodes the human ACE-2 protein of the invention.
 XX
 SQ Sequence 3396 BP; 1034 A; 659 C; 772 G; 931 T; 0 other;
 Alignment Scores:
 Pred. No.: 0 Length: 3396
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x ABR67623 (1-3396)
 QY 1 SerThrIleGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 136 TCCACCATGTGAGCAACGCGCCAGACATTTTGGACAGATTAAACCGAAGCGGAAGC 195
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheAsnThrAsnHisGluGluAsn 40
 DB 196 CTGCTTCATCAAGGTCACCTGCTTCGGAATTAAACCAATATTTACTGAAGCAAT 255
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerAlaPheLeuLysGluGlnSerThr 60
 DB 256 GTCCAAACATGATTAATGCTGGGAGCAATGCTGCTTTTAAAGGACAGTCCACA 315
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80

DB 316 CTTCGCCAATGATCCATCAACAAGAAATTCAGAAATCTGCAGCTTCACGTCGAC 375
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 376 GCTCTTACGCAAAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 435
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 DB 436 ATCTTAATATCAATGAGCACCACATCTACAGTACAGTACAGTACAGTACAGTACAGTAC 495
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 496 CAAGAATGCTTATTAATGAGCAGGTTGGAATGAATGAATGAATGAATGAATGAATGAAT 555
 QY 141 AsnGluArgGluLeuProAlaTyrPheGluSerTyrPheSerGluValGlyLeuGlnLeuArgPro 160
 DB 556 AATGAGAGGCTGCTGGCTGGGAAACCTGAGATCTGAGCTGGCAAGAGCTGAGGCA 615
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 616 TTATATGAGAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675
 QY 181 TyrGlyAspTyrTyrPheGlyAspTyrGluValAsnGlyValAspLysTyrAspTyrSer 200
 DB 676 TATGGGATTTATGGAG 735
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisTyrPheGluGluIleLysProLeuTyrGlu 220
 DB 736 CGGGCCAGTGTATTAATGAG 795
 QY 221 HisLeuHisAlaTyrValAlaArgAlaLysLeuMetAsnAlaTyrProSerTyrLysSerPro 240
 DB 796 CATCTTCATGCTATGTGAGGCAAGTGAATGAATGAATGAATGAATGAATGAATGAATGA 855
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheThrAsnLeu 260
 DB 856 ATGGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 915
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 916 TACTCTTACAGTTCCTTGGACAGAAACCAATGATGATGATGATGATGATGATGATGATGAT 975
 QY 281 AspGlnAlaTyrPheAspAlaGlnArgIlePheLysGluAlaGluLysPheValSerVal 300
 DB 976 GACACAGGCTGGAGTACAGAGAAATATCAAGAGAGGCGAAGTCTTGTATCTGTT 1035
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrPheGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 1036 GGTCTTCTTAATATGATCTCAAGGATCTGGGAAATTCATCTCAACGCGCCAGGAAT 1095
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrPheLeuGlyLysGlyAspPheArgTle 340
 DB 1096 GTTCAGAAAGAGCTGTCCCATCCACAGTGTGGACCTGGGGAAGGCGCATTCACAGATC 1155
 QY 341 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisHisGluMetCylHis 360
 DB 1156 CTTATGTGCAAAAGGTGCAATGAGCAGCTTCGACAGCTCATCAGATGAGGCGCAT 1215
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuAsnGlyValAsnGlu 380
 DB 1216 ATCCAGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
 QY 381 GlyPheHisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1276 GGATTCATGAAAGCTGTGGGAAATCATGTCATCTTCTGCTAAGAAATGGAGCTTAAATCAA 1335
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 DB 1336 AATCATCATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1395
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLys 440
 DB 1396 CTGCTCAACAGACCTACGATGTGTGGAGTCTGCCATTTACTTACATGTTAGAGAG 1455


```

QY 441 TTPATGTPMValPheLysGlyLileProLysAspGlnTPMLeLysTPTrp 460
DB 1456 TGGAGGTGATGCTTTAAAGCGGCAATTCACCAAGACAGTGGATGAAAAAGTGGCG 1515
QY 461 GIMuLysARGLiileValGlyValValGluProValProHisAspGluThrTyrCys 480
DB 1516 GAGATGAGGAGGATGATGTTGGGGTGGTGGACCTGTGGCCCATGATGAAACATACCTGT 1575
QY 481 ASPProAlaSerLeuPheHisValSerAsnAspTyrSerPheileArgTyrTyrThrArg 500
DB 1576 GACCCCGCATCTGTTCCATGTTTCATATGATTAATCATCATCATCATATACCAAG 1635
QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaAlaLysHisGluGlyPro 520
DB 1636 ACCCTTACCAATTCACATTCACATTCACATTCACATTCACATTCACATTCACATTC 1695
QY 521 LeuHisLysCysAspPileSerAsnSerThrsGluAlaGlyGlnLysLeuPheAsnMetLeu 540
DB 1696 CTGCACAAATGATGATCTCAAACTCTACAGAACCTGACAGAACTGTTCAATATCTGT 1755
QY 541 ArgLeuGlyLysSerGluProTyrPheLeuAlaLeuGluLysAsnValValGlyAlaLysAsn 560
DB 1756 AGCGTTGAAATCAGAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 1815
QY 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPheLysAspGln 580
DB 1816 ATGAAATGAAAGCCAGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 1875
QY 581 AsnLysAsnSerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIle 600
DB 1876 AACAAATCTCTTTGTTGGGATGAGTACGACGAGTACGAGTACGAGTACGAGTAC 1935
QY 601 LysValArgLysSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPheLysAsp 620
DB 1936 AAGGTGAGATTAAGCTTAAATACAGCTTGGAGATTAACATATGATTAAGTGAACAG 1995
QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLysVal 640
DB 1996 GAAATGATCTGTTCCGATCATCTGTTGATCATCTGATGAGGATCTTTTAAAGTA 2055
QY 641 LysAsnGlnMetLileuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
DB 2056 AAAAATCAGATGATCTTTTGGGAGGAGATGTCGAGTCTGATTTGAAACAGAG 2115
QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspLileIleProArgThr 680
DB 2116 ATCTCTTAATTTCTTTGTCATCTGACCTTAAATAATGTCTGATATCTATCTAGAACT 2175
QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
DB 2176 GAGCTTGAAAGGCGCATACAGAGATGTCGAGAGCCGATCATATGATCTTTCCGCTGAAAT 2235
QY 701 AspAsnSerLeuGlnPheLeuGlyLileGlnProThrLeuGlyProProAsnGlnProPro 720
DB 2236 GACACACAGCTTACAGATTTCTGGGATACAGCCAACTGTGACCTCTAACACAGCCCT 2295

```

```

XX Key Location/Qualifiers
FH 5'UTR 1..81
FT CDS /*tag- a
FT 82..2499
FT /*tag- b
FT /product= "Human ACE-2 protein"
FT sig_peptide 82..135
FT mat_peptide /*tag- c
FT 136..2496
FT /*tag- d
FT /product= "Human mature ACE-2 protein"
FT /EC_number= "3.4.15.1"
FT 136..2301
FT /*tag- e
FT /note= "Encodes extracellular domain"
FT 1201..1215
FT /*tag- f
FT /note= "Encodes zinc binding domain (ZBD)"
FT 2302..2376
FT /*tag- g
FT /note= "Encodes transmembrane domain"
FT replace (2239, G)
FT /*tag- h
FT /note= "Single chain polymorphism: This polymorphism
FT results in an amino acid change from Asn to Asp"
FT 2377..2496
FT /*tag- i
FT /note= "Encodes cytoplasmic domain"
FT 2500..3396
FT /*tag- j
FT replace (2834, T)
FT /*tag- k
FT /note= "Single chain polymorphism"
PN WO200212471-A2.
XX 14-FEB-2002.
PD 09-AUG-2001; 2001WO-US25059.
XX PF
XX 09-AUG-2000; 2000US-0635501.
PR (MILL-) MILLENNIUM PHARM INC.
PA
XX Action S, Robison KE, Hsieh FY;
PI WPI: 2002-257481/30.
XX P-PDB: AAE20353.
XX DR
XX Isolated human polypeptide, known as angiotensin converting enzyme-2,
XX useful for treating or preventing the development of an abnormal blood
XX pressure or related diseases, e.g. hypertension, heart failure or
XX myocardial infarction
XX
XX Claim 1; Fig 1; 21bp; English.
XX
XX The invention relates to human angiotensin converting enzyme-2 (ACE-2)
XX polypeptides and polynucleotides. ACE-2 is also known as peptidyl
XX dipeptidase A (EC 3.4.15.1). Polypeptides of the invention are useful
XX for treating or preventing the development of abnormal blood pressure
XX and diseases or disorders associated with the protein in a subject. The
XX diseases include hypertension, hypotension, congestive heart failure,
XX chronic heart failure, acute heart failure, myocardial infarction,
XX atherosclerosis, arrhythmia and renal failure. They are also useful
XX for treating inflammatory conditions and diseases relating to fertility.
XX The present sequence is human full-length ACE-2 cDNA. The ACE-2 gene is
XX located on the X chromosome at p21-22.
XX
XX Sequence 3396 BP; 1034 A; 659 C; 772 G; 931 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0 Length: 3396

```


Score: 3669.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AMD32586 (1-3396)

QY 1 SerThrIleGluGluGlnAlaValThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 136 TCACCATTCAGACAGGACGACATTTTGGACAAAGTTTAAACCAAGACGCGAGAC 195
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPasnTyrAsnThrAsnIleThrGluGluAsn 40
 DB 196 CTTTCTTATCAAAAGTTCACCTGCTTCTTGGAATTATTAACACCAATTTCTGAGAGAAAT 255
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrPserAlaPheLeuLysGluGlnSerThr 60
 DB 256 GTCCAAACATGAAATATGCTGGGACAAATGCTGCTTTTAAAGGACAGTCCACA 315
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 316 CTTCGCCAAATGATCCACTACACAGAAATTCAGAAATTCACAGTCAAGCTTCAGCTCAG 375
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 376 GCCTTCAGCAAAATGGCTCTTCAGTGCCTGTCAGAAAGACAAAGCAACGGTTGAGACACA 435
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnPro 120
 DB 436 ATTCTAAATACATGACACCACTACAGTACAGTGAAGAAAGTTGTGAACCCAGATTAATCCA 495
 QY 121 GlnGluCysLeuLeuLeuGluGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 496 CAGAAATGCTTATTACTTGAACCAAGTTTGAATGAATATGCAAAACAGTTTATGACTAC 555
 QY 141 AsnGlnArgLeuThrPalatrrpGluSerTrrpArgSerGluValGlyLysGlnLeuAspPro 160
 DB 556 AATGAGAGGCTCTGGCTGGGAAAGCTGAGAGATCTAGGTGGGCAAGCAAGCTGAGGCCA 615
 QY 161 LeuTyrGluGluTyrValIleLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 616 TTATATGAGAGATGTTGGTCTTAAATAATGAGATGCGCAAGCAAGCAATCTTTATGAGAC 675
 QY 181 TyrGlyAspTyrTrrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 DB 676 ATATGGGATATATGAGAGAGACTATGAAGTAAATGGGTAGATGCTATGACTACAGC 735
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 DB 736 CGCGGCGCATTTGATGAAGATGTGGAACATACCTTTAAGAGATTAACCATTAATATGAA 795
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 796 CATCTTCATGCCCTATGTCAGGCGCAAAAGTTGATGAATGCTTCTCTTATATACAGTCCA 855
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrrpGlyArgPheThrPheAsnLeu 260
 DB 856 ATGTGATGCCCTCCCTGCTCATTTGCTGGTGAATATGGGGGTAAATTTTGGACAAATCTG 915
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 916 TACTCTTTGACAGTTCCTTTGGACAGAAACCAACATAGATGTTACTGATGCAATAGTG 975
 QY 281 AspGlnAlaTrrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 976 GACCGAGGCTGGGATGACAGAGAAATATTTCAAGGAGCGGAGAAAGTTTGTATCTGT 1035
 QY 301 GlyLeuProAsnMetThrGlnGlyPheThrPglLysSerMetLeuThrAspProGlyAsn 320
 DB 1036 GGTCTTCTTAATATCTCAAGAGATTCGGGAAATTTCCATGCTAAACGACCCAGGAAT 1095
 QY 321 ValGlnLysAlaValLysHisProThrAlaTrrpAspLeuGlyLysGlyAspPheArgIle 340

DB 1096 GTTCAGAAAGCAGCTGCGCATCCACAGCTGGGACCTGGGAGGCGGACTTCAGATC 1155
 QY 341 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisHisGluMetGlyHis 360
 DB 1156 CTATATGTCACAAAGGTGACAAATGAGACACTCTCTGACAGCTATCATGATGAGGGCAT 1215
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1216 ATCCAGTATGATATGGCATATGTCACAAACCTTTTCTGTAAAGAAATGAGAGCTAAATGAA 1275
 QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerIleAlaIleProLysHisLeu 400
 DB 1276 GGATTCATGAAAGCTGTGGGAAATCATGCTCTTCTGACGCCACCACTTAACATTTA 1335
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnLysAspAsnGluThrGluIleAsnPhe 420
 DB 1336 AATCATATGGCTCTCTGTCACCCGATTTTCAAGAGACATGAACAGAAATTAACCTTC 1395
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1396 CTGCTCAAAACAGCACTACGATTTGTTGGACTGCTGCCATTTACTTACATGTGTAGAGAAC 1455
 QY 441 TrrpArgTrrpMetValPheLysGlyGluIleProLysAspGluTrrpMetLysTrrp 460
 DB 1456 TGGAGGTGATGCTTTTAAAGGGAAATTTCCAAAGCCAGTGTGATTAATAAGTGGTGG 1515
 QY 461 GluMetLysArgGluIleValGlyValGluProValProHisAspGluThrTyrCys 480
 DB 1516 GAGATGAAGCGAGATATGTTGGGTGGGAACTGCGCCATGATGAACAACTACTGT 1575
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
 DB 1576 GACCCCGCATCTCTCATGTTCTTAATGATTAATCATTCATTCGATATTAACAGAGG 1635
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 DB 1636 ACCCTTACCATTCAGTTTCAAGAGACCTTGTCAAGCGATTAACATGAAGGCCCT 1695
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
 DB 1696 CTGCACAAATGTGACATCTCAAACTCATAGAAAGTGGACAAACCTTCAATATGCTG 1755
 QY 541 ArgLeuGlyLysSerGluProTrrpThrLeuAlaLeuGluAsnValIleGlyAlaLysAsn 560
 DB 1756 AGGCTTGCAAAATCAGAACCTTCGACCCATGAGATGGAAGAAATGTTGAGAGCAAGAAC 1815
 QY 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrrpLeuLysAspGln 580
 DB 1816 ATGAATGTAAAGGCACCTGCTCAAACTTGTAGCCCTTATTAACCTGGCTGAAAGACAG 1875
 QY 581 AsnLysAsnSerPheValGlyTrrpSerThrAspTrrpSerProTyrAlaAspGlnSerIle 600
 DB 1876 AACMAAATCTTTTGTGGATGAGTACCGATGAGTCCATATGACACCAAAAGCAATC 1935
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrrpAsnAspAsn 620
 DB 1936 AAAGTGAAGGATTAACCTTAATACAGCTCTTGGAGATAAAGCATATGAATGAAGAACAT 1995
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
 DB 1996 GAATGTACCTGTTCCGATCATCTGTGATATGATATGAGGACAGACTTTTAAAGTA 2055
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
 DB 2056 AAAAATCAGATGATCTTTTGGGAGGAGATGTCGAGTGGCTTAATTTGAACCAAGCA 2115
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 DB 2116 ATCCCTTTAATTTCTTTGTCACCTGACCACTAAAGATGTGATATCATCTCTGAACT 2175
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700

Db 2176 GAAGTGAAGGCAATGAGATGTCGCCGAGCCGATATCATGATGCTTCCTCGTGAAT 2235
 QY 701 ASPASerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2236 GACAAACAGCCTAGAGTTCTGGGGATACAGCAACACTTGGACCTCTTAACAGACGCCCT 2295
 RESULT 7
 AAS14880
 ID AAS14880 standard; cDNA; 2920 BP.
 AC AAS14880;
 XX
 DT 20-DEC-2001 (first entry)
 XX
 DE Human cDNA encoding novel human protein NHP #1.
 XX
 KW Human; novel human protein; NHP; ss; antidiabetic; antirheumatic;
 KW antiarthritic; cytoskeletal; antiarteriosclerotic; valvulopathy;
 KW neuroprotective; neurotrophic; antiparkinsonian;
 KW anti-human immunodeficiency virus; antiasthmatic; vasotropic; cardiant;
 KW hypotensive; anorectic; antileptin; neuroleptic; anticonvulsant;
 KW antianemic; immunosuppressive; cerebroprotective; antimicrobial;
 KW immunomodulatory; antibacterial; antiparasitic; thymimetic;
 KW immunomodulator; antiseborrheic; dermatological; vasoconstriction;
 KW gastrointestinal disorder; cardiovascular disorder; hypertension;
 KW coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;
 KW cachexia; male infertility; impotence; testicular cancer; lung tumour;
 KW hyperproliferative disorder; pulmonary system disorder;
 KW central nervous system disorder; bone disorder;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW Huntington's disease; schizophrenia; mania; dementia; paranoia;
 KW panic disorder; learning disability; amyotrophic lateral sclerosis;
 KW psychosocial; autism; sleep disorder; immune system disorders;
 KW Hashimoto's thyroiditis; musculo-skeletal system disorders;
 KW multiple sclerosis; ischemic brain injury; stroke; infectious disease;
 KW diabetes mellitus; immunological disorder; asthma; AIDS;
 KW acquired immunodeficient syndrome; leukaemia; rheumatoid arthritis;
 KW inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;
 KW neural system disorder; respiratory disorder; olfactory disorder;
 KW wound healing; chromosome X.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 213..2348
 FT /tag="a
 FT /product="NHP #1"
 FT /transl_except="(pos:867..869,aa:Xaa)
 FT /transl_except="(pos:930..932,aa:Xaa)
 FT /transl_except="(pos:1707..1709,aa:Xaa)
 FT /note="Xaa= Any amino acid"
 XX
 PN WO2001/4896-A1.
 XX
 PD 11-OCT-2001.
 XX
 PE 02-APR-2001; 2001WO-US10542.
 XX
 PR 03-APR-2000; 2000US-194118P.
 XX
 PR 29-SEP-2000; 2000US-236384P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Moore PA, Ni J, Soppet DR, Coleman TA, Gentz RL, Endress GA;
 PI Li Y, Dillon PJ;
 DR WPI; 2001-626394/72.
 DR P-PSDB; AAO09092.
 XX
 PT New human proteins, useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cardiovascular
 PT disorders, autoimmune disorders and reproductive disorders
 XX

PS Claim 1; Page 291-292; 318pp; English.
 XX
 CC The invention relates to novel human proteins (NHP) and the
 CC nucleic acids that encode them and antibodies raised against them.
 CC The proteins, antibodies and nucleic acids are useful in the diagnosis,
 CC prognosis, prevention and/or treatment of diseases and/or disorders
 CC involving vasodilation, gastrointestinal disorders, cardiovascular
 CC disorders (e.g. hypertension, erectile dysfunction, high blood pressure,
 CC coronary heart disease and arteriosclerosis), anorexia, obesity, bulimia,
 CC cachexia, disorders of small intestine, disorders of reproductive system
 CC (e.g. male infertility and/or impotence), testicular cancer, lung tumours
 CC and other hyperproliferative disorders, disorders of pulmonary system,
 CC central nervous system disorders, bone disorders, neurodegenerative
 CC diseases and behavioural disorders (e.g. Alzheimer's disease, Parkinson's
 CC disease, Huntington's disease, schizophrenia, mania, dementia, paranoia,
 CC panic disorder, learning disabilities, amyotrophic lateral sclerosis,
 CC psychoses, autism, sleep disorders), immune system disorders (e.g.
 CC Hashimoto's thyroiditis), renal and musculo-skeletal system disorders,
 CC central nervous system disorders (e.g. multiple sclerosis, ischemic
 CC brain injury and/or stroke), infectious diseases, diabetes mellitus,
 CC immunological disorders (e.g. asthma, acquired immunodeficient syndrome
 CC (AIDS), leukaemia, rheumatoid arthritis, inflammatory bowel disease,
 CC sepsis, acne, psoriasis and lupus erythematosus), neural system
 CC disorders, respiratory disorders, olfactory disorders and wound
 CC healing. The present sequence encodes an NHP of the invention and
 CC is located on the X chromosome.
 CC
 SQ Sequence 2920 BP; 897 A; 568 C; 654 G; 788 T; 13 other;
 Alignment Scores:
 Pred. No.: 0 Length: 2920
 Score: 3833.00 Matches: 715
 Percent Similarity: 99.31% Conservative: 4
 Best Local Similarity: 99.17% Mismatch: 1
 Query Match: 22 Indels: 0
 DB: gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AAS14880 (1-2920)
 QY 1 SerThrIleGluGlnGlnAlaIleThrPheLeuAspLysPheAsnHisGluAlaGluAs 20
 Db 83 TTCACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 142
 QY 20 PLeuPheTygIleSerSerLeuAlaSerTPAsnTyAsnThAsnIleThrgluGluAs 40
 Db 143 CCTGTTCTATCAAGTTCACCTTCTCTGGAATATTAACCCAAATTTACTGAAGGAA 202
 QY 40 nValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerTh 60
 Db 203 TGTCCAAACATGATATATGCTGGGGACAAATGCTGCTTTTAAAGAACAGTCCAC 262
 QY 60 rLeuAlaGlnMetTygPheLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuG 80
 Db 263 ACTTGCCCAATGATATCCTACACAGAAATTCAGAACTCAGCTTACGCTGCA 322
 QY 80 nAlaLeuGlnGlnAsnGlySerValLeuSerGluAspLysSerLysArgLeuAsnTh 100
 Db 323 GGCCTCTTCAGCAAAATGGGCTTCAGTCTCAGAGACAGAGCAAAAGGTTGAACAC 382
 QY 100 rIleLeuAsnThrMetSerThrIleTygSerThGlyLysValCysAsnProAspAsnPr 120
 Db 383 AATCTTAATACATGAGCAGACCATCTACAGTACGAGAAAGTTGTAAACCAATATATCC 442
 QY 120 oGlnGlnCysLeuLeuLeuGlnIleProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTy 140
 Db 443 ACAAGAAATGCTTATTAATTGACACAGGTTGAATGAATATATGCAAAAGTTAGACTA 502
 QY 140 rAsnGluArgLeuTrpAlaTrpGluSerTPAsnSerGluValGlyLysGlnLeuArgPr 160
 Db 503 CAATGAGAGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGGTGGGCAAGCAGCTGAGGCC 562
 QY 160 oLeuTygGlnGluIleTygValValLeuLysAsnGluMetAlaArgAlaAsnHisTygGlu 180

Db 563 ATTATATGAGAGATGTTGCTTGAATAATGAGATGCAAGCAATCATTTATGAGA 622
 QY 180 PTYGLYASPTYTPATPGIYASPTYRGLUVALASNGLYVALASPGLYTYRASPTYSE 200
 Db 623 CTATGGGATATATGGAGAGAGACTATGAATGAGGATGAGATGCTATGACTACAG 662
 QY 200 FARGGLYGLNLEULIEGLIUSAPVALIUSIETHRPHGLUGLULIEYSPROLEUTYGI 220
 Db 683 CCGGGCGCAAGTTGATGATGAGATGGAACAATACCTTTGAAGAGATTAAACATATATGA 742
 QY 220 UHISLEUHSIALATYVALATAGALYLSLEUMETASALATYRPROSERIYRILLESER 240
 Db 743 ACACCTTCATGCTCTATGTGAGGCCAAGTTGATGAAGCCATTCCTTCATATACATCC 802
 QY 240 OILGLCYLSLEUPROALHISLEULEUGLYASPMETRPGIYARPHETRPHASNE 260
 Db 803 AATGGATGCTCCCTGCTGCTATTTGCTGGGATATGTGGGATGATTTGGACAAATYT 862
 QY 260 UTYRSELEUTHRALPROPHGLIYGLNLYSPROASNILEASPYALTHRASPALAMEVA 280
 Db 863 GTACMSFTTGACAGTTCCTTGAGACAGAAACCAACATAGATGTACTGATGCAATGCT 922
 QY 280 LASGLNALATRPASPLAGINARGIIEPHELISGLIUALAGLUYSPHEVALISERVA 300
 Db 923 GGACCAAGCTGGGATGACAGAAATATTCAGAGAGCCGAGAGATCTTTGATCTGT 982
 QY 300 IGLYLEUPROASMETRTHINGIYPHETRPGIUSASERMETLEUTHRASPROGLIYAS 320
 Db 983 TGGCTTCTCATATATGATCAGAGATTCCTGGGAAATTCATGCTAACGGACCCAGAAA 1042
 QY 320 NVALISLALVALALYLSHISPROTHRALATRPASPLEUGLYLSGLYASPHETARGI 340
 Db 1043 TGTTCAGAAACAGCTGTGCCATCCCAAGCTTGGGACCTGGGAGGCGCATCTTCAGAT 1102
 QY 340 ELEUMETCYSHRILYVALIETHMETASAPSPHELEUTHRALHISGLIUMETGIYH 360
 Db 1103 CCTATGTGCAAAAGGTGACAAATGAGAGACTTCTGACAGCTATATGATGAGGGCA 1162
 QY 360 SILGLINTYRSPMETALATYRALAALAGINPROPHLEULEUTHRAGSNGLYALAASNG 380
 Db 1163 TATCAAGATGATATGAGATATGCTGCACAAACCTTTCTGCTAAGAAATGAGACTATATGA 1222
 QY 380 UGLYPHENISGLUALVALIGLYLIULEMETERLEUSERALALATHRPROLYSHISLE 400
 Db 1223 AGGATTCATGATACCTGTGGGAAATCATGCTCTTGCACACCACTAAGCAATTT 1282
 QY 400 ULYSERTILEGLYLEULEUSERPROASPHENGLIUSAPSNILUNHRGIULIENAPH 420
 Db 1283 AAAATCCATGTGCTTCTGCTACCCGATTTTCAAGAAACATGAACAAATTAACCTT 1342
 QY 420 ELEULEULYGLNALALEUTHRILEVALIGLYTHRLEUPROPHETHRTRYMETLEUGI 440
 Db 1343 CCTGCTCAAAACAGACATCAGATGTGGGACTCTGCATTTACTTACATGATTAAGAAA 1402
 QY 440 STRPATRTMETVALPHELYSGLYLIULEPROLYSAPGLINTRPHETLYSTPTPT 460
 Db 1403 GTGGAGGTGATGCTTTAAAGGGGAAATCCCAAGACCAAGGATGAAGAAAGGTGTG 1462
 QY 460 PGLUMETYSARGIULILEVALIGLYVALIULIULIULIULIULIULIULIULIULI 480
 Db 1463 GGAATGTGAAGCAGATAGTTGGGGGTGTGGAACTGTGCCCATGATGAACATATCTG 1522
 QY 480 SASPROALASERLEUPHENISVALISERASNPYRSEPHETLEARGTYRTHRAR 500
 Db 1523 TGACCCGCACTCTGCTCATGTTTCTAATGATTAATGATTAATGATTAATGATTAATG 1582
 QY 500 GTHRLEUTYRGLNPHENGLNPHENGLIUALALEUCYSGIUALALALYSHISGLIYPT 520
 Db 1583 GACCTTTACCAATTCAGTTTCAAGAGCACTTTGTCACACCACTTAACATGAAGGCC 1642
 QY 520 OLEUHSIYSCYASPLIESERASNSERTHRGLIUALAGLYLINYSLEUPHEANMETLE 540
 Db 1643 TCTGCAAAATGTGATCTCAAACTTACAGAAAGCTGGACAGAAAGCTGTTCAATATGCT 1702

QY 540 UARGLEUGLYLSSESGIUPROTRPHRLEUALALEUGIUSANVALIGLYALALYAS 560
 Db 1703 GAGGNTTGGAAATATCAGAAACCTTGACCTGACATTTGGAATAATGTGAGAGCAAA 1762
 QY 560 NMEASVALARPROLEULEUASNTYRPHGLIUPROLEUPHETHTRPLEULYASPGI 580
 Db 1763 CATGATGTAGAGCCAGCTGCTCACTACTTTAGGCCCTTATTTACTGCTGGAAGACA 1822
 QY 580 NASNLYSASERSEPHVALIGLYTRPSERTHRASPTRPSERPROYRILASPGINSE 600
 Db 1823 GACCAAGATTTCTTTTGGGATGAGTACAGCTGAGTCCATATGCAAGCAAGCAT 1882
 QY 600 ELYSVALARGIIESELEULYSESERALALEUGLYASPYALATYRGIUTRPSANSPAS 620
 Db 1883 CAAAGTGAAGATAGCCTAAATACAGCTCTGTGAGATTAAGCATATGATGAAGACAA 1942
 QY 620 NGLUMETIYRLEUPHEARGSERSEVALALATYRALAMETARGINTYRPHLEULIYVA 640
 Db 1943 TGAAGTGTACCTGTTCCGATCATCTGTTCATATCTATGAGGCGATCTTTTAAAGT 2002
 QY 640 ILYSASNGIUMETILEUPHEGLYGLIUSAPVALARGVALALASANLEULYSPROAR 660
 Db 2003 AAAAATCGATGATTTCTTTTGGGAGAGAGATGTGCGATGCTAATTTGAACCAAG 2062
 QY 660 GILSESPHEASNPHEVALIETHRALPROLYSASNVALISERAPILIELEPROARGTH 680
 Db 2063 AATCTCTTAATTTCTTTGTCACATGACCTAAATAATGTGCTGATATCATCTCTAGAAC 2122
 QY 680 RGLUVALIGLYSALALILARGMETSERARGERTGLIEASAPALAPHEARGLEUAS 700
 Db 2123 TGAAGTTGAAAGGCGCATAGATGTCCGGAGCGGTACCAATGATGTCTCCGTGAA 2182
 QY 700 NASPASERLEUGIUPHEGLIYGLINPROTHRLEUGIYPROPPOASNGIUPROPR 720
 Db 2183 TGACGACACCTTAGATTTCTGGGATRACGCCAACACTGTGACCTCTTACACAGCCCC 2242
 QY 720 720 2243
 Db 2243 T 2243

RESULT 8
 AAS14890 standard; cDNA; 2911 BP.
 AAS14890:
 20-DEC-2001 (first entry)
 Human cDNA encoding novel human protein NHP #11.
 Human; novel human protein; NHP; ss; antidiabetic; antirheumatic;
 antiarthritic; cytoskeletal; antiarteriosclerotic; vulnery;
 neuroprotective; nootropic; antiparkinsonian;
 anti-human immunodeficiency virus; antiasthmatic; vasotropic; cardiac;
 hypotensive; anorectic; antilethargic; neuroleptic; anticonvulsant;
 antianemic; immunosuppressive; cerebroprotective; antimicrobial;
 antiinflammatory; antibacterial; antiparasitic; thyromimetic;
 immunomodulatory; antiseborrheic; dermatological; vasoconstriction;
 gastrointestinal disorder; cardiovascular disorder; hypertension;
 coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;
 cachexia; male infertility; impotence; testicular cancer; lung tumour;
 hyperproliferative disorder; pulmonary system disorder;
 central nervous system disorder; bone disorder;
 neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 Huntington's disease; schizophrenia; mania; dementia; paraneoplasia;
 panic disorder; learning disability; amyotrophic lateral sclerosis;
 psychosis; autism; sleep disorder; immune system disorder;
 Hashimoto's thyroiditis; musculo-skeletal system disorders;
 multiple sclerosis; ischemic brain injury; stroke; infectious disease;
 diabetes mellitus; immunological disorder; asthma; AIDS;
 acquired immunodeficient syndrome; leukaemia; rheumatoid arthritis;
 inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;

neural system disorder; respiratory disorder; olfactory disorder;
wound healing.

Homo sapiens.

Key Location/Qualifiers
CDS 213..998
FT /tag= a
FT /product= "NHP #11"

MO200174896-A1.

11-OCT-2001.

02-APR-2001; 2001KW-US10542.

03-APR-2000; 2000US-194118P.
29-SEP-2000; 2000US-236384P.

(HUMA-) HUMAN GENOME SCI INC.

Moore PA, Ni J, Soppet DR, Coleman TA, Gentz RL, Endress GA;
Li Y, Dillon PJ;

WPI: 2001-626394/72.
P-PSDB; AAD09102.

New human proteins, useful for diagnosing, treating, preventing and/or
prognosing disorders related to the proteins, including cardiovascular
disorders, autoimmune disorders and reproductive disorders -
Claim 1; Page 297-298; 318pp; English.

The invention relates to novel human proteins (NHP) and the
nucleic acids that encode them and antibodies raised against them.
The proteins, antibodies and nucleic acids are useful in the diagnosis,
prognosis, prevention and/or treatment of diseases and/or disorders,
involving vasoconstriction, gastrointestinal disorders, cardiovascular
disorders (e.g. hypertension, erectile dysfunction, high blood pressure,
coronary heart disease and arteriosclerosis), anorexia, obesity, bulimia,
cachexia, disorders of small intestine, disorders of reproductive system
(e.g. male infertility and/or impotence), testicular cancer, lung tumours
and other hyperproliferative disorders, disorders of pulmonary system,
central nervous system disorders, bone disorders, neurodegenerative
diseases and behavioural disorders (e.g. Alzheimer's disease, Parkinson's
disease, Huntington's disease, schizophrenia, mania, dementia, paranoia,
panic disorder, learning disabilities), immune system disorders (e.g.
psychoses, autism, sleep disorders), amyotrophic lateral sclerosis,
Hashimoto's thyroiditis), renal and musculo-skeletal system disorders,
central nervous system disorders (e.g. multiple sclerosis, ischemic
brain injury and/or stroke), infectious diseases, diabetes mellitus,
immunological disorders (e.g. asthma, acquired immunodeficient syndrome
(AIDS), leukaemia, rheumatoid arthritis, inflammatory bowel disease,
sepsis, acne, psoriasis and lupus erythematosus), neural system
disorders, respiratory disorders, olfactory disorders and wound
healing. The present sequence encodes an NHP of the invention.

Sequence 2911 BP; 896 A; 570 C; 655 G; 788 T; 2 other;

Alignment Scores:

Pred. No.: 0 Length: 2911
Score: 3785.00 Matches: 715
Percent Similarity: 99.31% Conservative: 1
Best Local Similarity: 99.17% Mismatches: 4
Query Match: 97.83% Indels: 3
DB: 22 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AAS14890 (1-2911)

QY 1 SerThrIleGluGluAlaAlaIysThrPheLeu-rAspLysPheAsnHisGluAlaGluAs 20
DB 83 TCACACATTGAGGAGCAAGCCAAAGCATTTTGGGACAAATTTAAACCAAGAGCCGAGA 142

QY 20 pLeuPheTYrGlnSerSerLeuAlaSerTrpAsnTrpAsnIleThrGluGluAs 40
DB 143 CCTGTTCTATCAAAAGTTCCTGCTTGGAAATTTAAACACCAATTTATGGAAGAGA 202
QY 40 nValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerTh 60
DB 203 TGTCAAAACATGATTAATGCTGGGGGCAAAATGCTGCTTTTAAAGAACATCTCAC 262
QY 60 rLeuAlaGlnMetTYrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuG 80
DB 263 ACTTGCCCAAAATGATACACTACCAAGAAATTCAGAAATCTCACAGTCAAGCTTCAC 322
QY 80 nAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnTh 100
DB 323 GGCTCTTCAGCAAAATAGGCTCTTCAGTCTCTTCAGAGACAGCAAGCAAAAGGTTA 382
QY 100 rIleLeuAsnThrMetSerThrIleTYrSerThrGlyLysValCysAsnProAspAsnPr 120
DB 383 AATCTTAAATACAAATGAGCAGCAGCATCTACAGTCTGGAAGTTGTAAACCCAGATATCC 442
QY 120 oGlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTY 140
DB 443 ACAAGATGCTTATTAATCTTGAACCAAGTTTGAATGAATTAATGCAACAGTTTGA 502
QY 140 rAsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPr 160
DB 503 CAATGAGAGCGCTCTGGGGCTTGGGAAAGCTGAGATCTGAGCTGGCAAGCAGCTGAG 562
QY 160 oLeuTYrGluGluTYrValValLeuLysAsnGluMetAlaArgAlaAsnHisTYrGluAs 180
DB 563 ATTATATGAAAGTATGTGCTTGAATAAATGATGCAAGACAAATCATTTATGAGGA 622
QY 180 pTYrGlyAspTYrTrpArgGlyAspTYrGluValAsnGlyValAspGlyTYrAspTYrSe 200
DB 623 CTATGGGATTTATGAGAGGAGCTATGAAAGTAAATGGGTATGAGCTATGACTCAG 682
QY 200 rArgGlyGlnLeuIleGluAspValGlnHisThrPheGluGluIleLysProLeuTYrG 220
DB 683 CCGCGCGAGTGTGATGGAAGATGTGGAACATACCTTGAACAGATTAACCATATATGA 742
QY 220 uHisLeuHisAlaTYrValAlaGluAlaLysLeuMetAsnAlaTYrProSerTYrIleSerPr 240
DB 743 ACATTTATGCTATGTGAGGCCAAGATTGATGAATGCCATTCCTCTATACATGATCC 802
QY 240 oIleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpAsnLe 260
DB 803 AATGAGATGCTCCTCGCTCATTTGCTTGGTATGATGTGGGTATGATTTGGACAAATCT 862
QY 260 uTYrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVa 280
DB 863 GTACTCTTGTGACAGTCCCTTGGACAGAAACCAACATATGATCTGATGCAATGCT 922
QY 280 lAspGlnAlaTrpAspAlaGlnArgIlePheLysGlnAlaGluLysPhePheValSerVa 300
DB 923 GGACGAGCGCTGGAGTGCAGAGAAATTTAAAGAGCGGAGAAAGTCTT-GRATGTGT 981
QY 300 lGlyLeuProAsnMetThrGlnGlyPheTYrGluAsnSerMetLeuThrAspProGlyAs 320
DB 982 TGTCTTCTCTTAATGACTCAAGGATTCGTGGAAATTTCACTGATACGAGACCCAGAAA 1041
QY 320 nValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgI 340
DB 1042 TGTTCAGAAACCAATGCTGCACATCCACAGCTTGGAGCCTGGGAGGAGGCACTTCAGAT 1101
QY 340 eLeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyH 360
DB 1102 CCTTATGTGCAAAAGGAGCAATGAGACGACTTCTACAGCTCATATGATGATGAGGGGCA 1161
QY 360 sIleGlnTYrAspMetAlaTYrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnG 380
DB 1162 TATCCAGATGATATGCAATGCTGCACAAACCTTTCTGCTAAGAAATGAGAGCTATATGA 1221
QY 380 uGlyPheHisGlnLysAlaValGlyIleMetSerLeuSerAlaAlaThrProLysHisLe 400

```

Db      1222 AGGATTCACGAGCTGTTGGGAAATCTGTCACCTTTCGACGACACCTTAACCATTT 1281
Qy      400 uylsSerIleGlyLeuLeuSerProAspPheGlnIuAspangluThrGluIleAsnPh 420
Db      1282 AAAATCCATGGCTCTCTGTCACCCGATTTTCAAGAACATGAACAGAAATTAACCTT 1341
Qy      420 eleuLeuGlnIleAlaLeuThrIleValGlyThrLeuProPheThrMetLeuGluY 440
Db      1342 CCGTCCTAAACAGACCTACAGATTTGGAGCTGCGCATTTACTACATGTTGAGAA 1401
Qy      440 STRPATGTPMetValPheIleGlyGluIleProIuAspGlnTrpMetIleYsTrp 460
Db      1402 GTGGAGGTGATGCTCTTAAAGGGGAAATTCACCAAGACAGTGGATGAAAAAGGTG 1461
Qy      460 pGIuMetIuAspGlnIleValGlyValGluProValProHIsAspGluThrTrpY 480
Db      1462 GGAGATGAAGCGAGATAGTTGGGTGGGAACTGGCCCATGATGAAGAAACATCTG 1521
Qy      480 SAAPProAlaSerLeuPheHIsValSerAsnAspTrpSerPheIleArgTrpThrAr 500
Db      1522 TGACCCCGCATCTCTGTCATTTCTATGATTTCTATTCATTCATTCGATATTACCAAG 1581
Qy      500 gThrLeuTrpGlnPheGlnPheGlnIuAlaLeuGlyGlnAlaIleYsHIsGluIYr 520
Db      1582 GACCCCTTACCAATTCAGATTTCAGAGAACACTTTGTCAAGCAGCTAAACATGAAGCCCC 1641
Qy      520 oIeuHIsYsCyAspIleSerAsnSerThrGluIuAgIuGlnIuLeuPheAsnMetIe 540
Db      1642 TCTGCACAAATGGACATCTC -AACTCTACAGAACGCGGACAAACTGTCTCAATTCCT 1700
Qy      540 uArgIeuGlyYsSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaYsAs 560
Db      1701 GAGGNTTGAAGAAATCAAGAACCCGAGACCTAGATTGGAAGAAATGTTGAGACAAAGAA 1760
Qy      560 nMetAsnValArgProLeuLeuAsnTrpPheGluProLeuPheThrTrpLeuYsAspI 580
Db      1761 CATGAATGTAAGGCGACCTGCTCAACATCTTGAAGCCCTTAATTAACCTGCTGAAGACCA 1820
Qy      580 nAsnIuAsnSerPheValGlyTrpSerThrAspTrpSerProTrpAlaAspIleSerI 600
Db      1821 GAACACAAATTTCTTTTGGGATGGAGTACCGACTGGAGTCATATGACCAACCAAGCANT 1880
Qy      600 eIuYsValArgIleSerLeuYsSerAlaLeuGlyAspIuAlaTrpGluTrpAsnAspAs 620
Db      1881 CAAGGTGAGATAGCTTAATAATCAGCTTGGAGATGAAGCATATGAAATGAAGACGACA 1940
Qy      620 nGIuMetTrpLeuPheArgSerSerValAlaTrpAlaMetArgGluThrPheLeuYsVa 640
Db      1941 TGAATGTACCTGTCGATCANTGTCATATGCTATGATGAGGAGTCTTTTAAAGT 2000
Qy      640 lIuYsAsnGlnMetIleLeuPheGlyGluIuAspValArgValAlaAsnLeuYsProAr 660
Db      2001 AAAAATACAGATGATTTCTTTTGGGAGGAGGATGTCGAGTGGCTAATTTGAACCAAG 2060
Qy      660 gIleSerPheAsnPhePheValIthrAlaProIuYsAsnValSerAspIleIleProArGth 680
Db      2061 AATCTCCTTAATTTCTTTGTCATCTGACCTAAATAAGTGTCTGATTCATCTCTAGAAC 2120
Qy      680 rGIuValGluYsAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAs 700
Db      2121 TGAAGTTGAAGAGCCATCAGAGATGTCGCGAGCCGATATCATATGATCTTTCCGTGAA 2180
Qy      700 nAspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPr 720
Db      2181 TGAGACAGAGCTAGAGTTTCTGGGATGATACACCAACACTTGAGACTCTTAACGACCCCC 2240
Qy      720 o 720
Db      2241 T 2241

```

```

ID      AAS21279 standard; cDNA: 3732 BP.
XX
AC      AAS21279;
XX
AC      24-OCT-2001 (first entry)
XX
DE      Human cDNA sequence encoding for PRO1885 polypeptide.
XX
KW      Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW      breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW      cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW      adipocyte; A-peptide; factor VIIa; gene therapy; ss.
XX
OS      Homo sapiens.
XX
PN      W02010466-AZ.
XX
PD      07-JUN-2001.
XX
PF      01-DEC-2000; 2000MO-US32678.
XX
PR      01-DEC-1999; 99MO-US28301.
XX      01-DEC-1999; 99MO-US28634.
XX      02-DEC-1999; 99MO-US28551.
XX      02-DEC-1999; 99MO-US28564.
XX      02-DEC-1999; 99MO-US28565.
XX      09-DEC-1999; 99US-0170262.
XX      16-DEC-1999; 99MO-US30095.
XX      20-DEC-1999; 99MO-US30911.
XX      20-DEC-1999; 99MO-US30999.
XX      30-DEC-1999; 99MO-US31243.
XX      06-JAN-2000; 2000MO-US00227.
XX      06-JAN-2000; 2000MO-US00376.
XX      11-FEB-2000; 2000MO-US03565.
XX      18-FEB-2000; 2000MO-US04341.
XX      18-FEB-2000; 2000MO-US04342.
XX      22-FEB-2000; 2000MO-US04914.
XX      24-FEB-2000; 2000MO-US05004.
XX      01-MAR-2000; 2000MO-US05601.
XX      20-MAR-2000; 2000MO-US07377.
XX      21-MAR-2000; 2000MO-US08439.
XX      30-MAR-2000; 2000MO-US08439.
XX      17-MAY-2000; 2000MO-US13705.
XX      22-MAY-2000; 2000MO-US14042.
XX      30-MAY-2000; 2000MO-US14941.
XX      02-JUN-2000; 2000MO-US15264.
XX      10-NOV-2000; 2000MO-US30873.
XX
PA      (GENE ) GENENTECH INC.
XX
PI      Baker KP, Beresini M, Deforge L, Desnoyers L, Flivaroff E, Gao W;
PI      Geritsen ME, Goddard A, Godowski PJ, Gurney AU, Sherwood S;
PI      Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
XX      WPI: 2001-408281/43.
XX      P-FSDB: AATU2207.
XX
DR      Isolated, secretory and transmembrane PRO polypeptide used to detect
XX      other PRO polypeptides, link bioactive molecules to cells expressing
XX      PRO polypeptides, and detect the presence of mammalian tumours e.g.
XX      lung, breast, prostate, cervical
XX
XX      Claim 3; Fig 71; 813pp; English.
XX
XX      AAS21244-AAS21518 encode for novel human secretory and transmembrane
XX      PRO polypeptides. The PRO polypeptides are useful to detect other
XX      PRO polypeptides, to link bioactive molecules to cells expressing
XX      PRO polypeptides, to modulate biological activities of cells expressing
XX      PRO polypeptides, and to detect the presence of mammalian lung, colon,
XX      breast, prostate, rectal, cervical or liver tumours by comparing PRO
XX      polypeptide expression in a cell sample to that in a control sample.
XX      Some of the 275 sequences are also useful to stimulate the release of

```

CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes' antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 XX

SO Sequence 3732 BP; 1137 A; 722 C; 821 G; 1052 T; 0 other;

Alignment Scores:

Align. No.:	Length:	3732
Score:	Matches:	717
Percent Similarity:	Conservative:	1
Best Local Similarity:	Mismatches:	2
Query Match:	Indels:	138
DB:	Gaps:	1

US-09-978-385-2_COPY_19_738 (1-720) x AAS21279 (1-3732)

QY 1 SerThrTleGIuGLuGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 94 TCCACCACTTGAAGACAGGCGCAAGACATTTTGGACAAAGTTTAAACCAAGCGGCAAGAC 153
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheAsnThrAsnHisGluGluGluAsn 40
 DB 154 CTGTCTATCAAAAGTTTCACTTCTCTGGAATTAATACCACTTCAAGCTTCAAGGAGT 213
 QY 41 ValGlnAsnMetAsnAlaGlyAspLysTyrSerAlaPheLeuGluGlnSerThr 60
 DB 214 GTCCAAACATGCAATATGCTGGGACAAATGCTGCTTTTAAAGAACGTCACCA 273
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 274 CTGGCCCAATGATCCACTACAGAAATTCAGAACTCAGCTCAAGCTTCAGCTGCGAG 333
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysAsnThr 100
 DB 334 GCTCTCAGCAAAATGGCTTCAGCTTCAGCTTCAGCAAGCAAGCAAGCTTGAACACA 393
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyValCysAsnProAspAsnPro 120
 DB 394 ATTCTAAATACAAATGAGCAACCACTACAGTACGAAAGTTGTAAACCAAGTATATCA 453
 QY 121 GlnGluCysLeuLeuLeuGluProGluLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 454 CAAAGATGCTTATATTGACCAAGCTTGAATGAATGAATGAATGAATGAATGAATGA 513
 QY 141 AsnGluArgLeuTyrAlaTyrGluSerTyrPheSerGluValGlyGlnLeuArgPro 160
 DB 514 AATGAGAGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGTGGCGCAAGCAAGCTGAGGCA 573
 QY 161 LeuTyrGlnGluTyrValIleLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 574 TTTATGAAAGAGATGTGTGCTTGAATAATGAGATGCAAGCAAAATCATTAATGAGGAC 633
 QY 181 TyrGluAspTyrTyrArgLysAspTyrGluValAsnGluValAspGluTyrAspTyrSer 200
 DB 634 TATGGAGATTTATGGAGAGAGACTATGAAATGAATGAATGAATGAATGAATGAATGA 693
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGlnGluIleLysProLeuTyrGlu 220
 DB 694 CCGGGCCAGTGTATGAAATGTGAAACATACCTTGAAGAGATTAACCAATTAATATGAA 753
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 754 CATCTTCATGCTATGTAGGCGCAAAAGTTGATGAATGCTTATCTTCTATATCAAGTCA 813

QY 241 IleGluCysLeuProAlaHisLeuLeuGluAspMetTyrArgPheThrAsnLeu 260
 DB 814 ATTTGATGGCTCCCTCTCTATTTGCTGTGTATGTGGGGAGATTTTGGCAAAATCTG 873
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 874 TACTCTTTCAGAGTTCCTTGGACGAAACCAACATAGATGTTACTGATCAATGGTGTG 933
 QY 281 AspGlnAlaTyrAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 934 GACAGGCTGGGAGTCCACAGAAATTCAGAGAGGCGGCAAGATTTCTTGTATCTGTT 993
 QY 301 GlyLeuProAsnMetThrGlnLysPheTyrPgiLysSerMetLeuThrAspProGluAsn 320
 DB 994 GGTCTCTTATATATGACTCAAGATTTCTGGAAAATTCATCTTACGACCCAGGAAAT 1053
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrPheLeuGlyLysGlyAspPheArgIle 340
 DB 1054 GTTCAGAAACAGAGTCCCATCCACAGCTTGGGACCTGGGAGGCGCAACTTCAGAGATC 1113
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHis 360
 DB 1114 CTATATGTCACAAAGTGCACATGGACGACTTCTGACAGCTCATCATGATGAGTGGGAT 1173
 QY 361 IleGlnTyrAspMetAlaTyrAlaIleGlnProPheLeuArgAsnGluValAsnGlu 380
 DB 1174 ATCCAGTATGATATGATGATATCTCTCACAACTTTCTGCTAAGAAATGGAGCTATATA 1233
 QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1234 GAATTCACGAAGCGTGGGGAATTCATGTCATCTTTCGACGCCACCTAAGCATTTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAsnGluThrGluIleAsnPhe 420
 DB 1294 AATTCATTTGCTCTCTCTGACCCCACTTTTCAAGAACATGAACGAAATTAACCTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1354 CTGCTCAAAACAGCACTACGATGTGTGGACTGTGCATTTACTATGATTTAGAGAG 1413
 QY 441 TrpArgTyrMetValPheLysGlyGluIleProLysAspGlnTyrMetLysTyrTrp 460
 DB 1414 TGGAGGTGAGGTCTTTTAAAGGGGAATTCACAAAGCAAGTGGATGATGAAAGTGGTGG 1473
 QY 461 GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCys 480
 DB 1474 GAGATGAAGCGAGATATGATGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1533
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArg 500
 DB 1534 GACCCCGCATCTCTGTCATGTTCTGATGATTTACTCATTCATTCATTCATTCATTC 1593
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 DB 1594 ACCCTTACCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1653
 QY 521 LeuHisLysCysAspLysSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 537
 DB 1654 CTGCACAAATGTGACATCTCAAACTCTACAGAGAGCTGAGAAACGTTT- GTAAGAAAT 1712
 QY 537 ----- 537
 DB 1713 ACCCTCAAAATGTTGAACCTCTCTAGTATTCAGTATTCATTCATTCATTCATTCATTC 1772
 QY 537 ----- 537
 DB 1773 TATTTGATTTCTTTGTTCTTAAAGAAATTTTATGGCTCAAAATGCTCATTTTCAA 1832
 QY 537 ----- 537
 DB 1833 ACCAAACATTTAAATTTGTGTGTCAGAGAACCTAGAACATCAACAAATGGGTGGGCA 1892
 QY 537 ----- 537

PS Claim 10; Page 104-109; 125pp; English.

XX The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-converting enzyme is a zinc metalloproteinase that plays roles in blood pressure regulation and fertility. Zace2 can be expressed by standard recombinant methodology. Zace2 polypeptides are useful for treating an inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis), diseases associated with inflammation like arthritis and enterocolitis, as targets for identifying modulators of zinc protease activity, for screening or identifying new angiotensin-converting enzyme (ACE) inhibitors, and as a basis for rational drug design for inhibitory molecules. The nucleic acids can be used to detect the expression of a Zace2 gene in a biological sample, as probes for in vivo diagnosis and for detecting and localizing Zace2 gene expression in tissue samples, to determine whether a subject's chromosomes contain a mutation in the Zace2 gene, and to detect aberrations associated with the Zace2 locus. CC Inhibitors of ACE are used for treating hypertension of various conditions, including left ventricular systolic dysfunction, progressive renal impairment, scleroderma renal crisis, congestive heart failure due to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be used to treat infertility while Zace2 antagonists are used for inducing infertility. The present sequence represents a cDNA encoding the mouse Zace2-5 protein.

XX Sequence 2638 BP; 802 A; 556 C; 611 G; 669 T; 0 other;

Alignment Scores:

Score:	3.9e-313	Length:	2638
Percent Similarity:	3251.00	Matches:	598
Best Local Similarity:	89.86%	Conservative:	49
Query Match:	83.06%	Mismatches:	73
	84.03%	Indels:	0
	22	Gaps:	0

US-09-978-385-2_COPY_19_738 (1-720) x AAC84368 (1-2638)

QY 1 SerThrIleGluGluGlnAlaIleValThrPheLeuAspIysPheAsnHisGluAlaGluAsp 20
 DB 160 TCCCTCCACCGAGAAATGCCAAGACATTTTAAACATTTATATCGAGAGCTGAAGAC 219
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheAsnThrAsnIleThrGluGluAsn 40
 DB 220 CTGCTTATCAAACTTCACTTCTTGTGGAATTATATACATACTACTGAAAGAAAT 279
 QY 41 ValGlnAsnMetAsnAlaGlyAspIysTyrSerAlaPheLeuIysGluGlnSerThr 60
 DB 280 GCCCAAAAGATGAGTGGAGCTGCAGCAAAATGCTGCTTTATGAGAAACAGTCAAG 339
 QY 61 LeuAlaGlnMetIyrProLeuGlnGluIleGlnAsnLeuThrValIleIysGlnIleGln 80
 DB 340 ACTGCCCAAGTTCTCTCACTACAGAAATCCAGATCCGATCAATCAAGCCTCACTACAG 399
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspIysSerIysArgLeuAsnThr 100
 DB 400 GCCCTTCAGAAATGGGTCTTTCAGCACTGCAGACAGAAAGAAACAGTTGAACACA 459
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyIysValCysAsnProAspAsnPro 120
 DB 460 ATTTGGAACACCACTGACACCATTTTACAGTACAGTGAAGAGTTTGCACCAAGAACCA 519
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 520 CAAGATTCCTTATCTTGAAGCCAGATGGATGAATATGGCGAACAACACAGACTAC 579
 QY 141 AsnGluArgLeuTyrPalaTyrPgluSerTyrPargSerGluValGlyIysGlnLeuArgPro 160
 DB 580 AACCTGATGCTCTGGGATGGAGGCTGAGGAGCTGAGGTGGCAAGCAGCTGAGCGC 639
 QY 161 LeuTyrGluGluTyrValValLeuIysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 640 TTGTATGAAGAGTATGTGCTACTGAAAGACAGATGGCAAGAGCAACATTAATACAC 699
 QY 181 TyrGlyAspTyrTyrPargGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200

DB 700 TATGGGATTTATTTGAGAGGGGACTATACAGACAGAGCAAGATGGCTACACTATAC 759
 QY 201 ArgGlyGlnIleGlnAlaAspValGluIleThrPheGluGluIleIysProLeuTyrGlu 220
 DB 760 GGTAAACCTGATATGAAAGATGATACAGCTTCCGACAGATCAACGCCATTATGAG 819
 QY 221 HisLeuHisAlaTyrValArgAlaIysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 820 CATCTTCATGCTATGTGAGGAGAAAGTTGATGATACCTTCTTACATCAGGCC 879
 QY 241 IleGlyCysLeuProAlaHisLeuLeuIysAspMetIyrPgluArgPheThrAsnLeu 260
 DB 880 ACTGATGCTCTCCCTCCCATTTGCTGTGATGTGGGTGATATTTGACAAACCTG 939
 QY 261 TyrSerLeuThrValProPheGlyGlnIysProAsnIleAspValThrAspAlaMetVal 280
 DB 940 TACCTTTGACCTGCTTCCCTTTCACAGAAACCAACATAGATGTTACTGATGCAATGATG 999
 QY 281 AspGlnAlaTyrPalaGlnArgIlePheIysGluAlaGluIysPheValSerVal 300
 DB 1000 AATCAGGGCTGGAGTGCAGAAAGATATTTCAAGAGCAGCAAAATCTTGTTCCTGTT 1059
 QY 301 GlyLeuProAsnMetThrGlnGlyPheThrPgluAsnSerMetLeuThrAspProGlyAsn 320
 DB 1060 GGCTCTTCATATGATGATCAAGATCTGGCAAACTGTATGCTGACTGACCCAGCAT 1119
 QY 321 ValGlnIysAlaValCysHisProThrAlaTyrAspLeuGlyIysGlyAspPheArgIle 340
 DB 1120 GGCCGGAAAGTTGTCTGCCACCCACAGCTTGGATGTGGGACAGAGACTTGAGATC 1179
 QY 341 LeuMetCysThrIysValThrMetAspPheLeuThrAlaHisIleGluMetGlyHis 360
 DB 1180 AAGATGTGTCAAAAGTGCATGAGCAACATCTTGTGACAGCCATCAGAGATGGGACAC 1239
 QY 361 IleGlnTyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsnGlyValAsnGlu 380
 DB 1240 ATCCATATGACATGACATGCAATGCCAGCACTTCTCTGTAAGAAAGGACCAATGAA 1299
 QY 381 GlyPheHisGluAlaValGlyIleMetSerLeuSerAlaAlaThrProIysHisLeu 400
 DB 1300 GGGTTCATGAAAGCTGTGGAAGAAATCATGCTACTTGTGCGAGTACCACCAAGCATCTG 1359
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnIleAspGluThrGluIleAsnIle 420
 DB 1360 AAATCCATGCTGCTTCTGCAATCCGATTTTCAAGAAAGATGGAAGCAAGATTAACCTC 1419
 QY 421 LeuLeuIysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluIys 440
 DB 1420 CTATCTAAGAACAGCATGACATTTGTGGAACATACCTGTTACTTACTATGATGAGAG 1479
 QY 441 TyrArgTyrMetValPheIysGlyIleIleProIysAspGlnTyrMetIysIysTyrTrp 460
 DB 1480 TGGAGTGTGATGTCTTCTTGGGGTGAATTTCCAAAGACAGATGATGAAAAGGCTGG 1539
 QY 461 GluMetIysArgGluIleValGlyValIleGluProValProHisAspGluThrTyrCys 480
 DB 1540 GAGATTAACCGGAGATGCTTGTGTGTGGAGGCTCTGCTCATGATGAACATACCTGT 1599
 QY 481 AspProAlaSerLeuPheHisIleValSerAsnAspTyrSerPheIleArgTyrThrArg 500
 DB 1600 GACCTGCAATCTGTTTCATGTTCTATGATTAATCACTCATTCATTCGATATTAACAAG 1659
 QY 501 ThrLeuTyrGlnPheIleIleGlnGluAlaLeuCysGlnAlaIleHisIleGluIlePro 520
 DB 1660 ACCATTACCAATTCAGATTCAGAAAGCTCTTTCAGAGCAGCAAGATTAATGCTCT 1719
 QY 521 LeuHisIysCysAspIleSerAsnSerThrGluAlaGlyIleIysLeuPheAsnMetLeu 540
 DB 1720 CTGCACAAATGTGACATCTCAAAATTCACATGGAAGCTGGGCAAAAGTGTCTCAAGTCTG 1779
 QY 541 ArgLeuGlyLysSerGluProTyrThrLeuAlaLeuGluAsnValValGlyValAlaIysAsn 560


```

Db      |||||||.....:|||||
640 TTGTATGAAAGATAGTGTCTCTGAAAAACAGAGAGCGCAAGCAAAACAAATTAACACAC 699
Qy      181 TYRGLYASPTYTRPARGLYASPTYGLIYVALASNGLYVALASPLGLTYASPTYSER 200
Db      700 TATGGGATATATGAGAGGAGACATATACAGAGGAGAGAGATGGCTACAACTATTAAC 759
Qy      201 ARGGLYGLILEUILEGLIUSPVAGLIUHSITRPHGLUGLIULEYSPROLEUTYGLU 220
Db      760 CGTAAACAGATGATGAGATGATGAGAACGATCTCCGACAGATCAACCACTGTATGAG 819
Qy      221 HISLEUHSIALATYVALARGALALYSLEUMETASNAIATYRPROSETTYLISERPRO 240
Db      820 CATCTTCATGCTATGATGAGAGAGAGATGATGATGATGATGATGATGATGATGATG 879
Qy      241 ILEGLYCYSLERPROALAHISLEUENGLIYASPMETTRPGIATARGPHETRPHASNU 260
Db      880 ACTGATGCTCTCCCGCCCATTTGCTGTGATATGTTGGGTAGATTTTGGACAAATCTG 939
Qy      261 TYRSERLEUTHRALPHROPHGLYGLIUSPROASNIIEASPVALTIRASPALAMEVAL 280
Db      940 TACCTTTGACTGCTCTTGCACAGAACCAACATAGATGATGATGATGATGATGATG 999
Qy      281 ASPGINALATRPASPALAGLNARGIIEPHGLYGLIUSGLIULAGLIULYSPHEPHEVAL 300
Db      1000 AATCGGGCTGGGATGCAAGAAAGATATTTCAGAGCGCAGAAATCTTTGTTCTGTT 1059
Qy      301 GLYLEUPROASMETTRHGLIYPHETRPGIUSANSEMETLEUTHRALSPROGLIYAS 320
Db      1060 GGCCTTCCTCATATGACATCAAGAGANTCGGGCAAACTATGCTGACGACGACGACGAT 1119
Qy      321 VALGINLYSALVALCYSHISPROTHRALATRPASPLEUGLYGLYASPHARGTLE 340
Db      1120 GGGCGAAAGTGTCTGCCACCCACACCTTGGGATCTGGGACGCGGAGCTTGAGATC 1179
Qy      341 LEUMETCYSTRHLYSVALTHRMETASPSAPHELEUTHRALAHISGLIUMETGLYHS 360
Db      1180 AAGATGTTACAAAGGTCAACATGACACACTTTTGACAGGCCATCAGAGATGGGACAC 1239
Qy      361 ILEGLIYRASPMEALATYRALALAGIINPROHELEUENGLIULASNGLIYALASGLU 380
Db      1240 ATCCAAATATGACATGCGATATGCCAGGCAACCTTCCGCTAAGAAACGAGCAAGAA 1299
Qy      381 GLYPHEHISGLIULAVAGLYGLIULEMETERLEUSERIALAIALTHPROLYSHISLEU 400
Db      1300 GGGTCCCATGAGCGTGTGGAATCTATGCTCTTCTGAGCTACCCCAACACATCTG 1359
Qy      401 LYSSEITIEGLIYLEULEUSERPROASPHGLINGIUSPASNGIULTHGLIULEASPH 420
Db      1360 AAATCCATTTGGTCTTCTCCATCCGATTTTCAAGAGATGACGAAACGAGATTAACCTC 1419
Qy      421 LEULEULYSGIINALALEUTHRIEVALIGLYTHLEUPROPHETHTYRMELEUGLIULYS 440
Db      1420 CTACTGAAACAGCATTCACAATTTGTTGGAACACTACCGTTACTTACATGTTAGAGA 1479
Qy      441 TRPARGTTRMETVALPHELYSGLYGLIULEROLYSASPLINTPMETLYSLYSTRTP 460
Db      1480 TGGAGGTGATGGCTTCTGGGGTGAATTCCTCAAGAGCATGATGAAAGTGGTGG 1539
Qy      461 GLIUMETLYSARGIULIIEVALIGLYVALIGIULPROVALPROHISASPLIUTHTYR 480
Db      1540 CAGATGAAGCGGAGATCTGGTGTGTGAGAGCCCTGCTGCTGATGAACATCTCTGT 1599
Qy      481 ASPPROALASERLEUPHEHISVALSERASNAAPTYSERPHETLEARGTYRTHARG 500
Db      1600 GACCGTCATCTGTTCCATGTTCTAATATATCATCATCATCATCATATTAACAAG 1659
Qy      501 THRLEUTYGLINPHEGLINPHEGLIULALEUCYSGIINALAIALYSHISGLUGLYPRO 520
Db      1660 ACCATTACCAATTCAGATTCAAGAGCTCTTGTCAAGAGCTAGATTAAGGTTCT 1719
Qy      521 LEUHSIUSCYSPASPILESERASNSERTHGLIULAGLIULYGLIUSLEUPHEASMETLEU 540

```

```

Db      1720 CTGCACAAATGTACATCTCAAAATTCACATGAGAGCTGGCGAGAGTTGCTCAAGATCTG 1779
Qy      541 ARGLEUGLYSSSERGLUPROTRPHLEUALALEUGLIUSANVALVALGLYALALYSASN 560
Db      1780 AGCTTGGAAATTCAGAGCCCTGACCGAAGCTTGGAAAATGTGTAGAGACAGAGAT 1839
Qy      561 METASNVALARGPROLEULEUASNTYRPHGLUPROLEUPHETHTRPHLEULYSASPLIN 580
Db      1840 ATGATGTAAACACACTGCTCATTTACTTCCAAACGCTGTGTGATGACTGGCTGAAGAGCAG 1899
Qy      581 ASNLYSANSERPHEVALIGLYTRPSERTHASPTRISERPROTYRALASPLINSEITIE 600
Db      1900 AACGAAATTTCTTTGTGGGGTGAACACTGAATGAGGCCATATATCCACCAAGACAT 1959
Qy      601 LYSVALARGIIESERLEULYSERVALALEUGLYASPLYALATYGLIULTRPASNASPAN 620
Db      1960 AAGTGAAGATTAAGCTTAATAATAGCTTTGGAGCTAATGATATATGACCAACACAC 2019
Qy      621 GLUMETRYLEUPHEARSGERSEVALALATYRALAMELARGIINTYRPHLEULYSVAL 640
Db      2020 GAATGTTCTGTTCCGATCATCTGTTCATATGCCATGAGAAAGATTTCTCAATATAC 2079
Qy      641 LYSANGLIMETLIELEUPHEGLYGLIUSPVALARGVALALASNULEUSPROARG 660
Db      2080 AAAACACAGACAGTCTCTTTCTAGAGAGAGATGATGATGATGATGATGATGATGATG 2139
Qy      661 ILESERPHEANPHEPHEVALTHRALAPOLYSASNAVALSERASPIELLEPROARGTHR 680
Db      2140 GTCTCTCTTACTTCTTTGTGACCTTCACCCCAAAATGTCTGATGATCTTCTAGAACT 2199
Qy      681 GLIUALIGLIUSALIAILEARGMETSERARGSERARGIIEASNPALAPHEARGLEUAN 700
Db      2200 GAAGTGAGAGATCCATCAGAGATGCTCGGGGCCCATCAATGATGCTTTGGCGTGAAT 2259
Qy      701 ASPASERLEUGLIUPHELEUGLYIIEGLINPROTHLEUGLIUPROBROSNGLINPRO 720
Db      2260 GATPACAGCTGAGATTTCTGGGGATCAACCCACACTTGACCCACTTACCAAGCTCTCT 2319

RESULT 13
ABK30270/c
ID ABK30270 standard; cDNA; 2350 BP.
XX
AC ABK30270;
XX
XX 23-APR-2002 (first entry)
XX
DE Human G-protein-coupled protease #40.
XX
KW Human; ss; gene; G-protein-coupled protease; gene therapy;
KW transgenic; protease mediated disorder; proliferative disorder;
KW differential disorder; developmental disorder;
KW hematopoietic disorder;
XX
OS Homo sapiens.
XX
PN US6331427-B1.
XX
PD 18-DEC-2001.
XX
PE 26-MAR-1999; 99US-0280116.
XX
PR 26-MAR-1999; 99US-0280116.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Robison KE;
XX
XX WPI, 2002-129545/17.
XX
DR New polynucleotides encoding protease homologs of the G-protein-coupled
XX PT protease family, useful in identifying agonists and antagonists for
XX PT diagnosis and treatment of protease mediated disorders -
XX

```


QY 508 GlnGluAlaLeuGys-GlnAlaAlaLysHisGluGlyProLeuHisLysCysAspIle 527
 DB 690 CAGAGACGACTTTTGTCCACACGCTAAACATGAAGGCCCTTGCACAAATGTGCATCTC 631
 QY 527 r-AsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGln 547
 DB 630 GTAACCTACAGAGAGCTGCAGAGAAAGCTTCATATGCTGAGGCTTGAAATACAGAAC 571
 QY 547 rTTPThrLeuAlaLeuGlnAsnValValGlyAlaLys-AsnMetAsnValArgProLeu 566
 DB 570 CCGAGACCCCTGACATGGAAATGTGTAGAGAGAGAAAGAAACATGAATGTAAGCCCTG 511
 QY 567 LeuAsn-TyrPheGlu-ProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 586
 DB 510 CTCACGACTCTTGAGCCCGCTTATTTACCTGCGTGAAGAACCAACAGAAATCTTTTG 451
 QY 586 ALGlyTrpSerThrAspTrpSerProTyrAlaAspGln-SerIleLysValArgIle-Se 605
 DB 450 TGGGATGGAGTACCCAGCTGAGTCCATATGACAGACCAAAACATCAAGTGAAGTAAAG 391
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPhe 625
 DB 390 CCTAAATCAGCTCTTGAGATTAAGCATTAATGATGACACATCAATGAATGTCCTGTT 331
 QY 625 eArgSerSerValAlaLysAlaMetArgGlnTyrPheLeuLysValLysAsnGln-MetI 645
 DB 330 CCGATCATCTGTGATATGCTATGAGCAGCTACTTTTAAAGTAAATCAAGCATGA 271
 QY 645 LeuPheGlyGluGln-AspValAlaLysAlaAsnLeuLysProArgLysSerPheAsn 664
 DB 270 TTTCTTTGGGAGAGAGATGTGAGTGTCTAATTTGAAACCAAGATCTCTTAT 211
 QY 665 PhePhe-ValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGlu 684
 DB 210 TTTCTTTGGTCACTGCCCTTAAAGTGTCTGATATCTTCTTGAAGTGAAGTTGAAA 151
 QY 684 sAlaIleArgMetSerArgSerArgLysAsnAspAlaPheArgLeuAsnAspAsnSerIle 704
 DB 150 GCCCATCAGATGTCCCGAGCCGATCATCAATGATGCTTCCGTCGATGACAAAGCCT 91
 QY 704 uGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 90 AGAGTTTCTGGGATACAGCAACACTTGACCTCTCTTAACAGCCCT 42
 RESULT 14
 AAC84367
 ID AAC84367 standard; DNA; 2415 BP.
 AC AAC84367;
 AC AAC84367;
 DT 19-MAR-2001 (first entry)
 DE Human Zace2 protein encoding degenerate sequence.
 KW Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
 KW zinc metalloproteinase; blood pressure; zinc protease; hypertension;
 KW ventricular systolic dysfunction; renal impairment; heart failure;
 KW scleroderma renal crisis; atherosclerosis; antiinflammatory; human;
 KW antihypertensive; bradykinin inactivator; ds.
 OS Homo sapiens.
 OS Homo sapiens.
 PN WO20070032-A1.
 PN WO20070032-A1.
 PD 23-NOV-2000.
 PD 23-NOV-2000.
 PF 03-MAY-2000; 2000WO-US11932.
 PF 03-MAY-2000; 2000WO-US11932.
 PR 13-MAY-1999; 99US-0311482.
 PR 27-AUG-1999; 99US-0384706.
 PR 27-AUG-1999; 99US-0384706.
 PA (ZYMO) ZYMOGENETICS INC.

XX piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
 XX WPL; 2001-025018/03.
 DR P-PSDB; AAB48095.
 XX
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 PT associated with inflammation such as arthritis and enterocolitis -
 PS Disclosure; Page 103-104; 125pp; English.
 XX
 CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood
 CC pressure regulation and fertility. Zace2 can be expressed by standard
 CC recombinant methodology. Zace2 polypeptides are useful for treating an
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
 CC diseases associated with inflammation like arthritis and enterocolitis,
 CC as targets for identifying modulators of zinc protease activity, for
 CC screening or identifying new angiotensin-converting enzyme (ACE)
 CC inhibitors, and as a basis for rational drug design for inhibitory
 CC molecules. The nucleic acids can be used to detect the expression of a
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and
 CC for detecting and localizing Zace2 gene expression in tissue samples,
 CC to determine whether a subject's chromosomes contain a mutation in the
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.
 CC Inhibitors of ACE are used for treating hypertension of various
 CC conditions, including left ventricular systolic dysfunction, progressive
 CC renal impairment, scleroderma renal crisis, congestive heart failure due
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
 CC used to treat infertility while Zace2 antagonists are used for inducing
 CC infertility. The present sequence represents a degenerate sequence
 CC encoding the human Zace2 protein.
 XX
 SQ Sequence 2415 BP; 494 A; 218 C; 398 G; 335 T; 970 other;
 Alignment Scores:
 QY 2 ThrIleGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGlnAspLeu 21
 DB 58 ACNATHGARGARCARCGNMAARACNTTYYTGAVTAARTTYAAYCAVARGCNGARAYTN 117
 QY 22 PheTyrGlnSerSerLeuAlaSerTTPAsnTyrAsnThrAsnIleThrGluGlnVal 41
 DB 118 TTYTAYCARWSNMSNTYTCNMSNTGGAATTAAYAAACNAAYATRAACNARGARAAYTN 177
 QY 42 GlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeu 61
 DB 178 CARAAVGAAYTAAYGCGNGAYAAARTGWSGCTTYYTNAARAKARASNACNTYN 237
 QY 62 ALAGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnGlnAla 81
 DB 238 GGNCAATGTATCCNTYNCARARATHCARAAAYTNACNGTNAARTNCARTNARGCN 297
 QY 82 LeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIle 101
 DB 298 YTNCAARCARAAYGCGNWSNGNTYTNWSNGAGAAARASNAARNGNTNAAYACNATH 357
 QY 102 LeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAsnProGln 121
 DB 358 YTNAAIACNATGWSNACNATHTAWSNACNGNAAGTNGTAATCCNGATAYCCNCRK 417
 QY 122 GluLysLeuLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsn 141
 DB 418 GARTGYTNTNTNTNGARCGNCGNTTAAYAGATATATGCGNAAYMSNTNGAYTAAYAY 477


```

QY 142 GIUAIGLEUTPRLAIGLTPGLUSETRPARISERGIUVALIGLYSGINLEUAARPROLEU 161
DB 478 GARMCNNTYTGCGCWTGGGARMSNTGGMNMGNGARNGINGMAARCAATYTMNGCNCNTN 537
QY 162 TYRGIUGIUTYRVALVALLEULYSANGLUMETALAAAGLAASNSHISTYRGLUASPTRY 181
DB 538 TAYGARARATAGTGTGTATYTAARAAAYGARATGGCMNGCNCNCAATYATYAGARATYAY 597
QY 182 GLASPTYTRPARGLIASPTRYGLIUALASNGIYVALASPGIYTYRASPTRYSERARG 201
DB 598 GGNMAYTAYTGTGMMGNGMGATYATGARCTMAAYGNGNGAYGNTATGATTAWSMNGN 657
QY 202 GLYGINLEULIEGLIASPVALIGLUISTHREGLUGIULIULIYSPROLEUTYRGLUHS 221
DB 658 GGNCAKYTNATHGARAGATNGARCAVACNTTYGARCARATHAARCCNYTWTAVGARGAY 717
QY 222 LEUHSALATYRVALARGALALASLEUMETASNLATYRPROSEPTYRILLESERPROILE 241
DB 718 YTNCAVGCNTAYGTINMGNCNAARNTNAATGAATCCNATYCCMSNTYATHTMSNCNATH 777
QY 242 GLYCYLEUPROALAHLSLEULEUGIYASPMETRPGIYARGPHEPTPTTHASNLLEUTYR 261
DB 778 GGNTRYCCNCCNCAATYTYTNGNGAYATGTGGGNGMNGNTYTGACNAATYTTAT 837
QY 262 SERLEUTHRALPHEPHEGLYGLINYSPROASNLLEASPYALITRASPALAMETVALASP 281
DB 838 WSNATNACNGTNCNTTYGGNCARACCAACNAAYTHAATGTCNCAATGTCNGAY 897
QY 282 GLUAAITRPAALAGIATARGIIEPHELYSGIULIAGIULYSPPHEPHEVALSERVALIGY 301
DB 898 CARCMTGGGAYGCNCARBMGNATYTAAGARCCNARARAKTYTITTYGTMSNGIUGN 957
QY 302 LEUPROASNMETIRGLINGIYEPHEPTGLUASNMETLEUTHRALPHEPHEGLIYASVAL 321
DB 958 YTNCAANATYAGACNCARGNTTYTGGGAARAAYSNATGYTNACNGAYCCNGMAAYGTN 1017
QY 322 GLINYSALAYALCYSHSPROTHRALATRPASPREULIYSGIYASPPHEPHEGLILEU 341
DB 1018 CARARCCNGNTGYCAYCCNACNGCWTGGATYTNNGNARAGNGAYTYTGMNATHYT 1077
QY 342 METCYSTHRLYSVALIHRMETASPPHELEUTHRALIASHLSGLUMETGLIYHSILE 361
DB 1078 ATGAYACNARAGTNACNATGAYGAYTYTNAACNCCAYCARGARATGGNCATAT 1137
QY 362 GLNTPASPMETALATYRVALAAGINPHEPHEULEUARYSNGIYLAASNGIUGLY 381
DB 1138 CARATYATATGGCNTAYGCNCACACCCNTTYTNTYTNMAAYGGNCNMAAYGARGGN 1197
QY 382 PHEHISGLIUALAYALIGIULIEMETSERLEUSERLALATHRPROLYSHISLEULYS 401
DB 1198 TTYCAYGARGCNGTNGNGARATHATGWSNYTMSNCCNCCNCAARCAAYTNMAR 1257
QY 402 SETIIEGLIYLEUSERPROASPPHEGLIULIASPASNGIULIULIETHLEASNPHELE 421
DB 1258 WSNATHEGANTNTNMSNCCNGAYTYTCARAGARAAAYCARACNARATHAATYTYTN 1317
QY 422 LEULYSGIUALALEUTHRLIEVALIGIYTHLEUPROPHETRYMETLEUGIULYSTRP 441
DB 1318 YTNARCARCANTNACNATHGTNGNACNYTCCNTTYACNTAYATGTYNARARTGG 1377
QY 442 ARGTRPMETVALPHELYSGIULIULIETPOLYASPGIULIETPMETLYSTYRTPRGLU 461
DB 1378 MGMTGATGTNTTYAARGNGARATHCCMAARGATCARITGATGARARARAGTGTGGAR 1437
QY 462 METLYSARGIULIETVALIGIYVALIULIUPROVALPROHISAPGLIUTHYRYSASP 481
DB 1438 ATGAAMGNGARATHTGTNGNGTNGARCCNGTCCNCAVAYCARACNTAYTGYGAY 1497
QY 482 PROALISERLEUPHEHISVALSERASNPTRYSERPHEIIEAGTYTYRTHARAGTTR 501
DB 1498 CCGCMMSNTYNTTYCAYGTNMSNAIYGATYATWSNTTYATHMGNTATYATACNMGNA 1557

```

```

QY 502 LEUTYRGINPHEGINPHEGINIUALALEUCYSGIUALALALYSHSIGIUGIYPROLEU 521
DB 1558 YTNATYCARITTYCARITTYCARAGACNTYNTGYCARCCNCCNARCAVARGAGCNCNTN 1617
QY 522 HISLYSCYASPILESERASNSERTHGLUALIGIULIYSPHEASNMETLEUARY 541
DB 1618 CAYARFYGATYNTHMSNAIYMSNACGAGCNGCNCARARATNTTYAATYATYTNMG 1677
QY 542 LEUGIYYSERGLUPROTPTPTTHLEUALALEUGIULASNYVALIGIYALATYASASMET 561
DB 1678 YTMGNARASNSNARCCNTGACNTYTCNGAYTNGARARAYGTNGTNGCNCNARARAY 1737
QY 562 ASNVALARGPROLEULEASNTYRPHIEGLUPROLEUPHETHTYRLEULYSAPGLIASN 581
DB 1738 AAYGTNMGNCNTNTNTNAATYATYTGACCCNTYTTACNCTGTYNARAGYCARAY 1797
QY 582 LYSASNSERPHEVALIGIYTPSERTRASPTRYSERPROTYRALASPCINSEIIELYS 601
DB 1798 AARAAYSNTTYGTNGTNGTNGTNGGACNAYTGSGMSNCNTAYCNCAYCARMSNATHAR 1857
QY 602 VALARGIIESEULEULYSSEERLALALEUGIYASPLYALATYGLIUTRPSANSPASNGI 621
DB 1858 GTMGNATHTMSNTYTNARMSNCGNTYNGGAYTAARGCAYTAYGARGAAYGAYAGAR 1917
QY 622 METYRLEUPHEARGSERVALATYRALAMETARGIIEPHEGLIYSPHEULYSVALYS 641
DB 1918 ATGTAYTNTYTNMGNSNMSNGTNGCNTAYGCATMGNCARATYTYTNAARGTNAAR 1977
QY 642 ASNGIMETIIELEUPHEGLIUGIULASPYALARGVALALASNULEULYSPROARGILE 661
DB 1978 AAYCARATGATHTYNTTYGNGARGARGAYGTMGNGTNGCNAAYTNAARCCNMGNAH 2037
QY 662 SERPHEASNPHEPHEVALIHRALAPROLYSASNYVALSERAPILIELEPROARGHGLU 681
DB 2038 WSNATYATATYTYTNTYTNACNCCNCAARAAAGTMSGATYATHTHCCNMGNCNGAR 2097
QY 682 VALIGIULYALALIEARGMETSERARGSERARGIIEASNPALAPHEARGLEUASNP 701
DB 2098 GTMGARAAACNATHMGNTGMSNMGMSNMGNAATHAAGATGNTYMGNTYNAAYGAY 2157
QY 702 ASNSEULEUGIUPHELEUGIIEGLINPROTHRLEUGIYPROPROASNGINPROPRO 720
DB 2158 AAYWSNTYATYTYTNGTNGTNGTNGTNGCNCNCAAYCARCCNCCN 2214

```

RESULT 15
 AAS42515
 ID AAS42515 standard; cDNA; 3474 BP.
 AC AAS42515;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human cDNA encoding an mdtt protein, clone LI:347572.1:2000MAY01.
 KW Human; molecules for disease detection and treatment; mdt; ss;
 KW Antiartherosclerotic; hepatotropic; antiproliferative; cyostatic;
 KW immunosuppressive; antidiabetic; antiasthmatic; neuroprotective;
 KW osteopathic; antilarthritic; cell proliferative disorder;
 KW arteriosclerosis; cirrhosis; poriasis; cancer; adenocarcinoma;
 KW leukaemia; breast cancer; autoimmune disorder; AIDS;
 KW acquired immunodeficiency syndrome; Addison's disease;
 KW diabetes mellitus; asthma; multiple sclerosis; osteoarthritis.
 XX
 OS Homo sapiens.
 XX
 PN WO200162922-A2.
 XX
 PD 30-APR-2001.
 XX
 PE 21-FEB-2001; 2001WO-050586.
 XX
 PR 24-FEB-2000; 2000US-0185213.
 PR 16-MAY-2000; 2000US-0205232.

PR 17-MAY-2000; 2000US-0205285.
 PR 17-MAY-2000; 2000US-0205286.
 PR 17-MAY-2000; 2000US-0205287.
 PR 17-MAY-2000; 2000US-0205323.
 PR 17-MAY-2000; 2000US-0205324.
 PA (INCYTE) INCYTE GENOMICS INC.
 PI Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;
 PI Chen A, D'Sa SA, Amshew S, Dahl CR, Dam TC, Daniels SE;
 PI Dufour GE, Flores V, Fong WT, Greenwalt LB, Hillman JK, Jones AL;
 PI Liu TF, Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A;
 PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;
 PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson SJ;
 DR WPI: 2001-570631/64.
 DR P-PSDB; AAU25463.
 XX
 PT New disease detection and treatment molecule polynucleotides and
 PT polypeptides, useful for diagnosis and treatment of arteriosclerosis,
 PT cirrhosis, psoriasis, cancer, autoimmune disorders, diabetes mellitus,
 PT asthma and multiple sclerosis
 PS Claim 1, Page 139-140; 183pp: English.
 XX
 CC The invention relates to novel human molecules for disease
 CC detection and treatment (mdt proteins) and the polynucleotides encoding
 CC them. The MDPY polynucleotides and polypeptides are useful for diagnostic
 CC and therapeutic purposes e.g. to diagnose and treat cell proliferative
 CC disorders (e.g. arteriosclerosis, cirrhosis and psoriasis) cancers (e.g.
 CC adenocarcinoma, leukemia and breast cancer) autoimmune disorders
 CC (e.g. acquired immunodeficiency syndrome (AIDS) and Addison's disease)
 CC diabetes mellitus, asthma, multiple sclerosis, osteoarthritis, and many
 CC more diseases given in the specification. The present sequence
 CC encodes an mdt protein of the invention.
 CC
 XX Sequence 3474 BP; 1019 A; 696 C; 781 G; 978 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 2.88e-271 Length: 3474
 Score: 2832.00 Matches: 650
 Percent Similarity: 85.66% Conservative: 19
 Best Local Similarity: 83.23% Mismatches: 49
 Query Match: 73.20% Indels: 63
 Gaps: 10
 US-09-978-385-2_COPY_19_738 (1-720) x AAS42515 (1-3474)
 QY 1 SerThrIleGluGluGlnAlaValSerPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 96 TCCACCATTCGAGCAAGCCCAAGACATTTTGGACAAAGTTTAAACGACGAAGCGAAGAC 155
 QY 21 LeuPheTyGlnSerSer-LeuAlaSer-TTpaSnyTyAsnThrAsnIleThGluGluA 40
 DB 156 CTGTCATCAAAAGTTCACGTTGCTCTTGGAATTTAAACCCAAATTTACTGAAAGA 215
 QY 40 snValGln-AsnMetAsn-AsnAlaGlyAspLys-----TTPSerAlaPheLeuLys 56
 DB 216 ATGTCCCAACACATGCAATTAAGTGTGGCGACAAATGTGTACGCCCTTTTAAAG 275
 QY 57 GluGlnSerThrLeuAlaGln-MetTyProLeuGlnGluIleGlnAsnLeuThrVal 76
 DB 276 GAAAGTTCACACCTTGGCCCAAGATGATCCACTCAA-----GCAAACTTCACGACATC 329
 QY 76 sleuGlnleuGlnAlaLeuGlnGlnGly-----SerSe 88
 DB 330 TCCACATGTCAACGCTTACGCTGTCACAGGCTTCTTCAAGCCATMAAAGTGTGAGTTC 389
 QY 88 rValLeuSer-----GluAspLysSerLys-ArgLeuAsnThrIleLeuAsnThrMet- 105
 DB 390 AGGTGGTCATTCACGACACCAAGCAACCGGTTTAAACACATTTCTTAATATACAA 449
 QY 106 -----SerThrIleTySerThrGlyLysValCysAsnProAspAsnProGlnGluC 123

DB 450 TGAAGCCACCAATCTTAACAGTAACGTGAAAGAACTGCTAAACCCAGATTAATCCACAGAT 509
 QY 123 yslLeuLeuGluUpProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyAsnGlu 143
 DB 510 GCTTATATCTTGAACACAGGTTGAATGAATTAATGCAAAACAGTTTGAAGTACAAATGAGA 569
 QY 143 rGluTPAlaITPGLuSerTTPArgSerGluValGlyGlnLeuAlaGProLeuTyG 163
 DB 570 GGCCTTGGGCTTGGGAAAGCTGGAAGCTGAGGTCGGCAAGCAGCTAGGCCATTAATG 629
 QY 163 lGluTyValValLeuLysAsnGluMetAlaAlaAlaAsnHisTyGluAsp-Tyr-Gl 182
 DB 630 AAGAGTATGTGCTTGAATAAATGAGTGGCAAGCAAAATCATTAAGAGACTATATGG 689
 QY 182 yAspTyTTPArgLysAspTyGluValAsnGlyValAsp---GlyTyAspTySerAr 201
 DB 690 GGATTAATGAGAGAGAGACTATGAAGTAAATGGGGTAAATAGTGAATATGTTACACCG 749
 QY 201 gGlyGlnLeuIleGluAspValGluHisThr-PheGluGluIleLysProLeuTyr-Glu 220
 DB 750 CGGCCAGTTGATGAAGATGTTGGAACATACCTGTTGAAGATTAATCAATTAATGAGAA 809
 QY 221 HisLeuHisAlaTyValArgAlaLysLeuMetAsnAlaTyProSerTyTleSerPro 240
 DB 810 CATCTTACGCCCTATGATGAGGCCCAAGTTGATGAATGCTTATCTTCTATATACATCA 869
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTTP-GlyArgPheTTPTrAsnLe 260
 DB 870 AATGATGCTCCCTCCGCTCATTTGCTGTGATATGTCGGGTGAGATTTGCAAAATCT 929
 QY 260 uTySerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVa 280
 DB 930 GATCTTTGACAGATTCCTCTTGGACAAACAAACATATGATGATGATGCAATGCT 989
 QY 280 lAspGlnAlaTTPAspAlaGlnArgIlePheLysGluAla-GluLysPheValSerV 300
 DB 990 GGACCAAGGCTGGATGACACAGAAATATTCAGAGATCCGCAAGACTCTTGTATCTG 1049
 QY 300 aIGlyLeuProAsnMetThrGlnGlyPheTTP-GluAsnSerMetLeu-TTPAspProG 319
 DB 1050 TTGGTCTTCTTATGATGACTAGAAATCTCGGCAAAATTCATGCTTATGAGCAACAGG 1109
 QY 319 yAsnValGlnLysAlaValCysHis-ProThrAlaTTPAspLeuGlyLysGlyAspPhe 339
 DB 1110 AATGTTCGAAAGCAAGCTGCAATCCCAACAGCTTGGACCTGGGAAGGCGCATCTCA 1169
 QY 339 rg-IleLeuMetCysThrLys-ValThrMetAspPheLeuThrAlaHisLysLys 358
 DB 1170 GAGATCTTATGTGCAAAAGGTTAAACAATGAGCAGATTCCTGACAGCTCATCATGAGAT 1229
 QY 358 tGlyHisIleGlnTyAspMetAlaTyAlaAlaGlnPro-PheLeuLeuArg-AsnGly 377
 DB 1230 GGGGCTATTCAGATATGATGATGCAATTCGCGGCAACTTTTCTGCTAAGGAATGGA 1289
 QY 378 Ala-AsnGluGlyPheHisGlnAlaValGlyLysIleMetSerLeuSerAlaAlaThrPr 397
 DB 1290 GCTTAATGAAGATTCATGAGAGCTGTGGGAAATCAATGCTTTCGACGCCACAC 1349
 QY 397 oLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGln---GluAspAsnGlu 416
 DB 1350 TAAAGCTTAAATTCATTTGCTTGTGACCCGAGTTTACAGCAAGCAATGAAAC 1409
 QY 416 rGluIleAsnPheLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTy 436
 DB 1410 AGAAATTAATCTCTGCTCAAAACACACTCAGATGTGGGACTCTCCATTAAGTAA 1469
 QY 436 rMetLeuGluLysTTPArgTTPMetValPheLys-GlyGluIleProLysAspGlnTTPM 456
 DB 1470 CATGTTAGAGAAATGAGGTGATGCTTAAACGGGAAATTCACCAAGACAGGTGGG 1529
 QY 456 eLysLys-TTPTrpGluMetLysArgGlu-IleValGlyVal-ValGluProValProH 475

Db 1530 TGAAGAGTGTGGGAGATGACCAAGCAATAGTTGGGTGTGTGGAACTGTGCCCC 1589
 QY 475 ISASPGIURHTYR-CysAspProAlaSerLeuPheHisValSerAsnSPTyrSerPhe 494
 Db 1590 ATGATGAACATATATCTGTGACCCCGCATCTCTGTCCATCTTCTATATGATCTATCT 1649
 QY 495 ILEAGTYRTHRRATGTHTRLEU-TyrGlnPheGlnPheGln-GluAlaLeu-CysGln 513
 Db 1650 ATTGGATATATACCAAGACCCGTGTAACCAATTCAGTTCAAGAAAGACCTTTGTCAA 1709
 QY 514 ALALALYSHISGLUGLYProLeuHisLys-CysAspIle-SerAsnSerThrGlu---A 532
 Db 1710 GCAGCTAAACATGAGGCCCTCTGCACAAATTTGTGACATTTCTAAATTTCTACGAACGTC 1769
 QY 532 IAGGLINLys-LeuPheAsnMetLeuArgLeuGlyLys-SerGluProThrLeuAl 551
 Db 1770 GTGGACAGACACTGTTCAATATGCTGAGGCTTGAAAACTCAGAACCCCTGAGCCCTAGC 1829
 QY 551 ALAUGLUAENValVal-GlyAlaAlaLysAsnMetAsnValArgPro-LeuLeuAsnTyrPhe 570
 Db 1830 ATTGGAAATGTTGTAGGACCAAGCAATGAAATGTAAGCCACCTGCTCAACTCTT 1889
 QY 571 GLUPROLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGlyTrpSerThr 590
 Db 1890 GACCCCTATTTACCTGCTGCGAAGACCAAGCAAAATTTCTTTGGGATGCACTAC 1949
 QY 591 ASPTSPSerProTyrAlaAspGlnSerIle-LysValArgIleSerLeuLysSerAla 610
 Db 1950 GACTGGAGTCCATATGACACCAAGCATCACAAGTAGAGATTAAGCTTAATACAGCTCT 2009
 QY 610 UGLY-AspLysAlaTyrGluTrpAsnAsp-AsnGluMetLysLeuPheArgSerSer-Val 629
 Db 2010 TGGCAGATTAAGATGATGAATGGAAGACCAAGCAATGTAAGCTGTTCCATATCTGCT 2069
 QY 629 LALAVALA---MetArgLysIlePheLeu-LysValLysAsnGlnMetIleLeuPheG 648
 Db 2070 TGGATATGTTTAATGAGGAGTACTTTTAAACAGTAAATAATTCAGATATCTTTTG 2129
 QY 648 IYGLUGLUAENValValAlaAsnLeuLysProArgIleSerPheAsnPheValT 668
 Db 2130 GGGGAGGAGATGTGCGAGTGGCTTAATTTGAAACCAAGAAATCTCTTAATTTCTTGCTCA 2189
 QY 668 HRALAPROLysAsnValSer-AspIleIleProArg-ThrGluValGluLysAlaIleArg 687
 Db 2190 CTGACCTAAATAATGTCTGTGATATCTATCTTAAAGAACTGAAAGGCAATCAG 2249
 QY 687 GMEISerArgSerArg-IleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheT 707
 Db 2250 GATGTCCCGGAGCGGCTACTGCAATGATGCTTCCGTGCAATGACAGACGCTAGAGTTTC 2309
 QY 707 EUGLYIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2310 TGGGGATACACCAACACTTGGACCTCTTAACCAAGCCCT 2350
 RESULT 16
 AAC84369
 ID AAC84369 standard; DNA: 2415 BP.
 AC AAC84369;
 XX
 XX 19-MAR-2001 (first entry)
 DE Mouse Zace2-5 protein encoding degenerate sequence.
 XX
 XX Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
 KW zinc metalloproteinase; blood pressure; zinc protease; hypertension;
 KW ventricular systolic dysfunction; renal impairment; heart failure;
 KW scleroderma renal crisis; atherosclerosis; antiinflammatory; mouse;
 KW antiatheritic; Bradykinin inactivator; ds.
 XX
 XX Mus sp.
 OS
 XX
 PN WO200070032-A1.

XX
 PD 23-NOV-2000.
 XX
 PF 03-MAY-2000; 2000MO-US11932.
 XX
 PR 13-MAY-1999; 99US-0311482.
 PR 27-AUG-1999; 99US-0384706.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
 XX
 DR WPI, 2001-025018/03.
 DR P-PSDB; AAB48097.
 XX
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 PT associated with inflammation such as arthritis and enterocolitis -
 XX
 PS Disclosure; Page 112-113; 125pp; English.
 XX
 CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood
 CC pressure regulation and fertility. Zace2 can be expressed by standard
 CC recombinant methodology. Zace2 polypeptides are useful for treating an
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
 CC diseases associated with inflammation like arthritis and enterocolitis,
 CC as targets for identifying modulators of zinc protease activity, for
 CC screening or identifying new angiotensin-converting enzyme (ACE)
 CC inhibitors, and as a basis for rational drug design for inhibitory
 CC molecules. The nucleic acids can be used to detect the expression of a
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and
 CC for detecting and localizing Zace2 gene expression in tissue samples,
 CC to determine whether a subject's chromosomes contain a mutation in the
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.
 CC Inhibitors of ACE are used for treating hypertension of various
 CC conditions, including left ventricular systolic dysfunction, progressive
 CC renal impairment, scleroderma renal crisis, congestive heart failure due
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
 CC used to treat infertility while Zace2 antagonists are used for inducing
 CC infertility. The present sequence represents a degenerate sequence
 CC encoding the mouse Zace2-5 protein.
 XX
 SQ Sequence 2415 BP; 488 A; 228 C; 397 G; 324 T; 978 other;
 Alignment Scores:
 Pred. No.: 2,19e-252 Length: 2415
 Score: 2640.00 Matches: 477
 Percent Similarity: 71.83% Conservative: 38
 Best Local Similarity: 66.53% Mismatches: 202
 Query Match: 68.23% Indels: 0
 DB: 22 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AAC84369 (1-2415)
 QY 4 GUGLUGLINALALYSThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPheTyr 23
 Db 64 GARGAARAAAGCNAARACNTTYTNAAYTAAYTAAYCARGARGCARGAAYTTNMSRAY 123
 QY 24 GINserSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsnValGlnAsn 43
 Db 124 CARMSNWSNYTNGCWNSTGGAAYTAAYTAAYACNAAYTAHACGARGAARAAAGCCNCARAR 183
 QY 44 MetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuGluGlnSerThrLeuAlaGln 63
 Db 184 ARGWSMGARGCNGCNCNARTRGWSNGCNTTYTAYGARGACARCAWSNAARACNCNCAR 243
 QY 64 MetTyrProLeuGluGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeuGln 83
 Db 244 WSNITYSNYTNCARARATHCARACNCCNATHATHAARMGNCARAYTNCARACNTNCAR 303
 QY 84 GlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeuAsn 103

QY 492 TyrSerPheIleArgTyrThrArgThrLeuTyrGlnPheGlnIleGlnIleLeu 511
 Db 1673 GTGCTTACATCAGATGATGTTGTCAGCTTCATCATCCAGTTCCACGAGGACG 1732
 QY 512 CysGlnAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 531
 Db 1733 TGGCAGGACCTGCTGACGACGACGACGACGACGACGACGACGACGACGACG 1792
 QY 532 AlG1GlnIlePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 551
 Db 1793 GCGGCGGACGCGCTGCGGACGCGGACGCGGACGCGGCTGCTGCTGCTGCTGCTG 1852
 QY 552 LeuGlnAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 571
 Db 1853 ATGACGCTGATCAGGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACG 1912
 QY 572 PheLeuPheThrPheLeuPheLeuPheLeuPheLeuPheLeuPheLeuPheLeu 588
 Db 1913 CCGCTGCTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1972
 QY 589 SerThrAspTyrPhePhePhePhePhePhePhePhePhePhePhePhePhePhe 599
 Db 1973 CAGTACACATGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACG 2005
 RESULT 18
 AAA38330
 ID AAA38330 standard; DNA; 4020 BP.
 AC
 XX AAA38330;
 DT 21-AUG-2000 (first entry)
 DE
 XX Human angiotensin-converting enzyme (ACE) coding region.
 XX
 KW Angiotensin-converting enzyme gene; ACE; coding region; polymorphism;
 KW polymorphic marker; cardiovascular disease; myocardial infarction;
 KW unstable angina; hypertension; atherosclerosis; stroke; prognosis;
 KW drug screening; treatment outcome; human; ds.
 KM
 XX
 OS Homo sapiens.
 PN MO200022166-A2.
 PD 20-APR-2000.
 XX
 XX 13-OCT-1999; 99WO-IB01678.
 XX
 XX 14-OCT-1998; 98US-0104286.
 PR 14-OCT-1998; 98US-0104302.
 XX
 PA (EURO-) EURONA MEDICAL AB.
 XX
 PI Norberg LT, Andersson MK, Lindstrom PHR, Jonsson L;
 XX WPI; 2000-318010/27.
 DR
 XX
 XX
 PT Assessing cardiovascular status in humans involves comparing test
 PT polymorphic pattern comprising polymorphic positions within genes
 PT encoding specific proteins, with reference polymorphic pattern
 PS
 PS Disclosure; Page 114-115; 126pp; English.

CC cardiovascular disorders such as myocardial infarction, unstable angina,
 CC hypertension, atherosclerosis and stroke. They are also useful for
 CC predicting the likely cardiovascular status of a patient given a
 CC treatment regimen comprising administration of cardiovascular drugs
 CC (e.g., ACE inhibitors, beta-adrenergic receptor antagonists (beta-
 CC blockers) or calcium channel blockers). One or more polymorphic markers
 CC provides a basis for predicting the outcome of a treatment regimen.
 CC Fragments of the genes comprising a polymorphic site may be used as
 CC primers and probes for detecting genetic polymorphisms or in molecular
 CC library arrays for high throughput screening. The genes, and the proteins
 CC they encode are useful in the screening of potential cardiovascular
 CC drugs. Determination of an individual's polymorphic pattern reduces or
 CC eliminates trial and error in selecting a treatment for a particular
 CC individual cardiovascular patient. It also provides the ability to
 CC eliminate patients from clinical trials who are predicted to be
 CC non-responsive, or at a risk for an adverse response, to a particular
 CC treatment regimen. Adverse results in an early trial can be evaluated to
 CC identify polymorphic patterns so that the adverse results can be
 CC correlated with a sub-population of the test population, permitting
 CC exclusion of such sub-populations from the treatment group. Beneficial
 CC drugs can be approved for use in the appropriate population, thereby
 CC decreasing the number of patients required for a clinical trial, which in
 CC turn decreases the duration and cost of such trials. Sequences A38328 and
 CC A38330 represent, respectively, intron 16 and the coding region of
 CC the human ACE gene (Genbank X62855, J04144). The polymorphic sites
 CC identified are 375A/C, 582C/T, 731A/G, 1060G/A, 1215C/T, 2193G/A,
 CC 2328A/G, 2741G/T, 3132C/T, 3387T/C, 3503G/C, 3906G/A; and a deletion of
 CC nucleotides 1451-1783 in intron 16.
 XX
 SQ Sequence 4020 BP; 857 A; 1261 C; 1174 G; 728 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,13e-122 Length: 4020
 Score: 1337.00 Matches: 255
 Percent Similarity: 61.05% Conservative: 118
 Best Local Similarity: 41.73% Mismatches: 204
 Query Match: 34.56% Indels: 34
 DB: 21 Gaps: 9
 US-09-978-385-2_COPY_19_738 (1-720) x AAA38330 (1-4020)
 QY 2 ThrIleGluGluGlnIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 21
 Db 1952 ACTGATGAGGCTGAGGCCAGCAAGTTGTGAGGAAATATGACCGCAATCCAGTGTG 2011
 QY 22 PheTyrGlnSerSerIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 39
 Db 2012 TGAACGAGTATGCCGAGGCCAATGGAATCACAACATCAGACAGACAGACG 2071
 QY 40 -----AsnValGlnAsnMetAsnAlaGlyAspLysTyrSerLeu 53
 Db 2072 AAGATTGCTGCACAGACATGCAATATGACCAACACAC----- 2113
 QY 54 PheLeuLysGluGlnSerThrLeuAlaGlnMetTyrProLeuGlnIleGlnAsnLeu 73
 Db 2114 -----CTGAAGTACGAGCACCAGGCCAGGAAGTTGATGTGACCAATTCGACAGAC 2167
 QY 74 ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGlu 93
 Db 2168 ACTATCAGCGGATCATTAAGAGGTTCAGGACCTAGAAAGCGGCGCTGCTGCCG 2227
 QY 94 LysSerLysArgLysAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLys 113
 Db 2228 GAGCTGAGAGGACTACACAGATCTGTTGGATATGAAACCACTACAGGCTGCCACT 2287
 QY 114 ValCysAspProAspAsnProGlnIleLysLeuLeuLeuGluProGlyLeuAsnGluLe 133
 Db 2288 GTGTGACACCGAATGCG-----AGCTGCTGCGAGCTCGAGCCAGATCTACAGAAATG 2341
 QY 134 MetAlaAsnSerLeuAspTyrAsnGluArgLeuTyrPalatyrPgluSerTyrArgSerGlu 153
 Db 2342 ATGGCCACATCCCGGAATATATGAGACCTGTATGAGCATGGAGGCTGCGAGACAG 2401

154 ValGlyLysGlnLeuArgProLeuTyrGluGluValLeuLysAsnGluMetAla 173
 2402 GCGGGAGAGCCATCTCCAGTTTACCCGAATACGTGGAATCATCAACAGAGCTGCC 2461
 174 ArgAlaAsnHisTyrGluAspTyrGluAspTyrGluAspTyrGluValAsnGly 193
 2462 CCGGTCAATAGGTATGTAGATGCAGGGAGCTGCGAGGTCTATGTACGAGACACCATCC 2521
 194 ValAspGlyTyrAspTyrSerArgGlyLeuLeuIleGluAspValGluHisThrPheGlu 213
 2522 CTGGAG-----CAAGACCTGGAGCGGCTCTCCAG 2551
 214 GluIleLysProLeuTyrGluHisLeuHisAlaTyrValAlaGluLysLeuMetAsnAla 233
 2552 GAGCGGAGCCACTTACTCTACCTGCACTGCTACGTCGCGCGGCGCCCTCACCGCTCAC 2611
 234 Tyr---ProSerTyrIleSerProIleGlyLysLeuProAlaHisLeuLeuGlyAspMet 252
 2612 TACGGGGCCCAACATCACTGAGGGGCCATCTGCTCCTCACCTGCTGGGAGACATG 2671
 253 TrpGlyArgPheThrPheAsnLeuTyrSerLeuThrValProPheGlyLysProAsn 272
 2672 TGGGGGAGACCTGCTCCACATCTATGACTTGGGTGGCCCTTCCCTTCAGCCCTCG 2731
 273 IleAspValThrAspAlaMetValAspGlnAlaThrAspAlaGlnArgIlePheLysGlu 292
 2732 ATGGACACCAAGAGGCTATCTAAAGAGGCTGGAGCCCGGAGAGATCTTTAAGAG 2791
 293 AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheThrPglAsn 312
 2792 GCTATGATTTCTTCCCTCCCTGGGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 312
 313 SerMetLeuThrAspProGlyAsnValGlnLysAlaValGlyHisPheThrPheAsp 332
 2852 TCGATGCTGGAGAGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2911
 333 LeuGlyLysGly---AspPheArgIleLeuMetCysThrLysValThrMetAspAspPhe 351
 2912 TTTCACACGCGAGGAGCTCCGATCAGACAGTGCACACCGGAGACTGGAGAGCTG 2971
 352 LeuThrAlaHisGlnMetGlyHisIleGlnTyrAspMetAlaTyrAlaAlaGlnPro 371
 2972 GTGGTGGCCCAACCAATGAGGCGCATCTCAGATTCATTCAGCAACAAGACTTACCT 3031
 372 PheLeuLeuArgAsnGlyAlaAsnGluGlyPheHisGlnAlaValGlyLysIleMetSer 391
 3032 GTGGCTTGGAGAGGAGTGCACACCGGCTTCCATGAGGCCATTTGGGAGCTGTAAGCC 3091
 392 LeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGln 411
 3092 CTCTCAGTGTCTACGCGCAAGCACTGCACAGTCTCACTGCTGAGAGTGGAGGCTGGC 3151
 412 GluAspAsnGluThrGluIleAsnPheLeuLysGlnAlaLeuThrIleValGlyThr 431
 3152 AGGAC---GAGCATGACATCAACTTCTGATGATGATGATGATGATGATGATGATGATG 3208
 432 LeuProPheThrTyrMetLeuGluLysTyrArgTyrMetValPheLysGlyLysIlePro 451
 3209 ATCCCTTACAGTACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3268
 452 LysAspGlnThrMetLysLysTyrPheGluMetLysArgGluIleValGlyValGlu 471
 3269 AAGGAAATATTAACCAAGAGTGGTGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 3328
 472 ProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSerAsnAsp 491
 3329 CCAAGTGGCCAGACATCAAGGTGACTTGGACCCAGGGGCCAATCTCCATCTCTCTCTG 3388
 492 TyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnAlaLeu 511
 3389 GTGCTTACATCAAGTACTTGTCTGAGCTTCAATCAACAGTCTCCAGAGGAGGAGGAGG 3448
 512 CysGlnAlaAlaLysHisGluGlyProLeuHisLysCysAspIleSerAsnSerThrGlu 531

3449 TCCAGGAGCTGGCCACAGCGGCCCTGACCAAGTGTGACATCTACCACTCAAGAG 3508
 532 AlaGlyLysLeuPheAsnMetLeuArgLeuGlyLysSerGluPheProThrLeuAla 551
 3509 GCCGGGAGCCCTGGGAGCCGATGAGTGGCTTCAATGAGTGGCTGAGGCGCGGAGGCC 3568
 552 LeuLysAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGlu 571
 3569 ATGCACTGATCAACGAGGCGCAAGCCCAACATGAGCGCTGAGCATGTGTGAGCTACTCAAG 3628
 572 ProLeuPheThrTyrPheLysAspGlnAsnLys-----AsnSerPheValGlyTyr--- 588
 3629 CCGGTGCTGAGCTGCTCTCCGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3688
 589 SerThrAspTyrSerProTyrAlaAspGlnSer 599
 3689 CAGTACAACTGAGAGCGGCAACTCCGCTGCTCA 3721
 Db
 RESULT 19
 AAK9395
 ID AAK9395 standard; DNA: 4022 BP.
 AC AAK9395;
 XX 27-JUN-2002 (first entry)
 Dt
 XX DNA of APP related human homologue hcp51674.
 DE
 XX Neuroprotective; neurotrophic; transgenic fly; Alzheimer's disease; Abeta;
 KW amyloid precursor protein; tissue-specific expression control; human APP;
 KW APP pathway modulator; gene therapy; gene; ds.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 148..3945
 FT /tag- a "protein of human homologue hcp51674"
 FT /product- "note- "No start codon"
 XX
 MO200226820-A2.
 PD 04-APR-2002.
 XX
 PF 01-OCT-2001; 2001MO-BP11345.
 XX
 PR 29-SEP-2000; 2000US-236893P.
 PR 14-JUN-2001; 2001US-298309P.
 XX
 PA (NOVS) NOVARTIS AG.
 PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 PI Cohen D, Dengler UJ, Finelli AL, Freuler F, Konsolaki M;
 PI Reinhardt MMH, Zisman S;
 XX
 DR WPI: 2002-315796/35.
 DR P-PSDB: AAO20501.
 XX
 PT New transgenic fly, containing DNA encoding an Abeta portion of human
 PT APP, useful for identifying agents which modulate the APP pathway and
 PT which can be used to treat Alzheimer's disease -
 PS
 PS Example 4; Page 93-94; 129pp; English.
 XX
 CC The invention relates to a transgenic fly whose genome comprises DNA
 CC encoding a polypeptide having the Abeta portion of human amyloid
 CC precursor protein (APP), fused to a signal sequence. The DNA sequence
 CC encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in
 CC the specification. The DNA sequence is operably linked to a tissue-
 CC specific expression control sequence. Expression of the sequence gives
 CC the fly an altered phenotype. The purpose of the invention is for
 CC identifying agents that inhibit or promote the expression and/or function

of genes or encoded polypeptides which modify the APP pathway. The agent is a compound, triple helix DNA, antisense oligonucleotide, double stranded RNA molecule, ribozyme, or particularly an antibody. It is used to treat conditions such as Alzheimer's disease. The agent can be used as an APP pathway modulator or in gene therapy. This polynucleotide sequence represents the DNA of the APP related human homologue hcp51674.

Sequence 4022 BP, 856 A, 1262 C, 1174 G, 728 T; 2 other:

Alignment Scores:

Pred. No.:	2,136-122	Length:	4022
Score:	1337.00	Matches:	255
Percent Similarity:	61.05%	Conservative:	118
Best Local Similarity:	41.73%	Mismatches:	204
Query Match:	34.56%	Indels:	3
	24	Gaps:	9

US-09-978-385-2_COPY_19_738 (1-720) x AA093935 (1-4022)

```

OY      2 ThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeu 21
DB      1954 ACTGATGAGCGCTGAGGCCAGCAAGTTTGTGGAGGAATATGACCGGACATCCAGGTGTG 2013
OY      22 PheTYrGlnSerSerLeuAlaSerTrpAsnTrpAsnTrpAsnTrpGlnGlu----- 39
DB      2014 TGGACGAGTATGCGCGAGCAACTGGAACATCAACACCAATCACCACAGACAGACAGC 2073
OY      40 -----AsnValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAla 53
DB      2074 AAGATTCTGCTGCAGAGAACATGCAATATGCAATATGCAACCCACACC----- 2115
OY      54 PheLeuLysGluGlnSerThrLeuAlaGlnMetTYrProLeuGlnGluIleGlnAsnLeu 73
DB      2116 -----CTGAACTAGACGGCACCCAGGACGAGAGTTGATGATGAAACATTCACAGACACC 2169
OY      74 ThrValIleGlnGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAsp 93
DB      2170 ACATCATAGCGGATCTTAAGAAAGAGTTCAGACCTAGAAACGGGACCTGCTGCCAG 2229
OY      94 LysSerIleAsnTrpLeuAsnTrpIleLeuAsnTrpMetSerThrIleTYrSerThrGlyLys 113
DB      2230 GACCTGAGAGAGATACAAAGATCCCTTGATATGGAACACACCTACAGCGTGGCCACT 2289
OY      114 ValCysAsnProAspAsnProGlnGluCysLeuLeuLeuGlnProGlyLeuAsnGluIle 133
DB      2290 GTGGCCACCCGGAATGCG-----ACCTGCTGCACCTGACCCGAGTCTGCGAATGTG 2343
OY      134 MetAlaAsnSerLeuAspTYrAsnGlnAsnGlnLeuTrpAlaTrpGlnSerTrpArgSerGlu 153
DB      2344 ATGGCCACGTCGCCGGAATATATGAACCTGTATGGGCAATGGAGGCGTGGGAGACAAG 2403
OY      154 ValGlyLysGlnLeuAsnProLeuTYrGluGlnTYrValIleLeuLysAsnGluMetAla 173
DB      2404 GCGGGAGAGACCATCCCGAGTTTACCCGAAATATGTAAGTCTGCACTCAACAGGCTGCC 2463
OY      174 ArgAlaAsnHisTYrGlnAspTYrGlnAspTYrArgIleAspTYrArgIleValAsnGly 193
DB      2464 CGGCTCATGAGCTATGTAGATGAGAGGAGGACTCTGAGAGGTCTATGTACAGAACACCATCC 2523
OY      194 ValAspGlyTYrAspTYrSerArgGlnLeuIleGluAspValGlnHisThrPheGlu 213
DB      2524 CTGGAG-----CAAGACCTCGAGGCGGCTCTCCAG 2553
OY      214 GluIleLysProLeuTYrGlnHisLeuHisAlaTYrValArgAlaLysLeuMetAsnAla 233
DB      2554 GAGCTGACGACCTCTACCTCAACCTGACATGCTACGCGCGGCGCCGACACCTGAC 2613
OY      234 Tyr---ProSerTYrIleSerProIleGlyCysLeuProAlaHisLeuLeuGlyAspMet 252
DB      2614 TACGGGGCCGACGACATCAACCTGGAGGGGCCCATTCCTGCTCACTGCTGGGGAGACATG 2673
OY      253 TrpGlyArgPheTrpThrAsnLeuTYrSerLeuThrValProPheGlnGlnLysProAsn 272
DB      253 TrpGlyArgPheTrpThrAsnLeuTYrSerLeuThrValProPheGlnGlnLysProAsn 272

```

```

DB      2674 TGGGGCAGACCTGCTCCACATCTATGAGTGTGGCTGCTCTCCCTTACGCCCCCTGC 2733
OY      273 IleAspValThrAspAlaMetValAspGlnAlaTrpAspAlaGlnArgIlePheLysGlu 292
DB      2734 ATGGACACCAACAGAGCTATCTTAACAGAGGCTGAGGCGCCAGAGAGATGTTAAAGAG 2793
OY      293 AlaGlnLysPhePheValSerValGlyLeuProAsnMetThrGlnGlnGlyPheTrpGluAsn 312
DB      2794 GCTGATGATTTCTTCACTCCCTGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2853
OY      313 SerMetLeuTrpAsnProGlyAsnValGlnLysAlaValCysHisProThrAlaTrpAsp 332
DB      2854 TCGATGCTGAGAAAGACACCGGCGGCGGAGGTGTCTGCTGCTGCTGCTGCTGCTGCTG 2913
OY      333 LeuGlyLysGly---AspPheArgIleLeuMetCysThrLysValThrMetAspAspPhe 351
DB      2914 TTCTTCAACAGGAGACTTCGCGATCAAGAGTGCACACCGGAGTCTGAGAGACTG 2973
OY      352 LeuThrAlaHisHisGluMetGlyHisIleGlnTYrAspMetAlaTYrAlaAlaGlnPro 371
DB      2974 GTGGTGGCCACACAGAAATGGGCGACATCCAGTATTCATGATGATCAAAAGACTTACT 3033
OY      372 PheLeuLeuArgAsnGlyAlaAsnGlnGlyPheHisGluAlaValGlyGluIleMetSer 391
DB      3034 GTGGCTTACAGGAGGAGGCGCCACACCGGCTTCATGAGGCGCATTTGGGAGCGTCTAGCC 3093
OY      392 LeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGln 411
DB      3094 CTCTCAGTGTCTTACCGCCCAAGCAGCTGCACAGTCTCAACCTTGTGACACAGTGGGCTGC 3153
OY      412 GluAspAsnGlnThrGluIleAsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThr 431
DB      3154 AGCGAC---GAGCATGACATCAACTTCTGATGAAATGCGCCTTGACAGAGTGCCTT 3210
OY      432 LeuProPheThrTYrMetLeuGlnLysTrpArgTrpMetValPheLysGlyGluIlePro 451
DB      3211 ATCCCTCAGTACCTCCTGCATCAGTGGCGCTGAGGAGTATTTGATGAAAGCATCAC 3270
OY      452 LysAspGlnTrpMetLysTrpTrpArgIleMetLysArgGlnIleValGlyValIleGlu 471
DB      3271 AAGGACACTATACACAGAGATGTGTGAGACCTCAGGCTGAAATGACAGGCGCTGCGCC 3330
OY      472 ProValProHisAspGlnThrTYrCysAspProAlaSerLeuPheHisValSerAsnAsp 491
DB      3331 CGAGTGGCCAGGACCTCAAGATGACTTGTACCCAGGCGCCAAAGTCCACATTCCTTAC 3390
OY      492 TyrSerPheLeuArgTYrTYrThrArgThrLeuTYrGlnPheGlnPheGlnGluAlaLeu 511
DB      3391 GTGCTTACATCAGTACTTGTTCAGCTTCACTCATCATCCAGTCTCACTTCCAGAGGACTG 3450
OY      512 CysGlnAlaAlaLysHisGlnGluProLeuHisLysCysAspIleSerAsnSerThrGlu 531
DB      3451 TGGCAGGACACTGGCCACACAGGCGCCCTGCAACAAGTGTGATCTACCACTCCCAAGAG 3510
OY      532 AlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGlnProTrpThrLeuAla 551
DB      3511 GCGGGGAGCGCTGCGAGCCGACCGCATGAAGCTGGGCTTCAAGAGGCGCTGGCGGAGACC 3570
OY      552 LeuGlnAsnValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTYrPheGlu 571
DB      3571 ATGCACTGATTCACAGGCGCCACCAATGATGAGCGCTGGCGCATTTGAGCTACTCAAG 3630
OY      572 ProLeuPheThrTrpLeuLysAspGlnAsnLys-----AsnSerPheValGlyTrp--- 588
DB      3631 CGGCTGCTGAGACTGGCTCCGACAGAGAAACGAGACTGCATGAGGAGAGAGCTGGGCGG 3690
OY      589 SerThrAspTrpSerProTYrAlaAspGlnSer 599
DB      3691 CAGTACAACTGGAGCGCAACTCCTGCTCA 3723

```

RESULT 20
AA004027
ID AA004027 standard; cDNA; 4024 BP.

Db 3209 ATCCCTTCAGCTACCTGTCGATCATGCGGCTGAGGATATTGATGAAGCATCACC 3268
QY 452 LysAspGlnTrpMetLysIleTyrPrlLumetLysAspGlnIleValValGlu 471
Db 3269 AAGGAACTATACCAAGAGAGGTGGAGCTCAGGCTGAAAGTACCAAGGCTCTGCCCC 3328
QY 472 ProValProHisAspGluTyrTyrCysAspProAlaSerLeuPheHisValSerAsnAsp 491
Db 3329 CCAGTCCCGCAGGACATCAAGAGTACTTGCACCGAGGCGCAGATTCCACATCTCTTACGC 3388
QY 492 TyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnAlaLeu 511
Db 3389 GTGCCTTACATCAGGTACTTGTCTCAGCTTCATCATCATCAGTTCCAGTCCAGGACACTG 3448
QY 512 CysGlnAlaAlaLysHisGluGlyProLeuHisLysCysAspIleSerAsnSerThrGlu 531
Db 3449 TCCCGAGCAGCTGGCCACACGGGCCCCCTGCACAAAGTGTGACATCTACCACTCCAGAG 3508
QY 532 AlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProTyrPrlLeuAla 551
Db 3509 GCCGGGACAGCGCTGGCGACCGCATGAAGCTGGGCTTCAGTAGGCGCTGGCCGGAAGCC 3568
QY 552 LeuGlnAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGlu 571
Db 3569 ATGCAGCTGATCACGGGCGACGCCCAATGAGCCGCTCGGCCCATGTGAGCTACTTCAAG 3628
QY 572 ProLeuPheThrTyrPrlLeuLysAspGlnAsnLys-----AsnSerPheValGlyTyr--- 588
Db 3629 CCGCTGCTGAGACTGCTCGCCAGAGAAAGAGCTGATGGGAGAAAGCTGGGCTGGCCG 3688
QY 589 SerThrAspTyrSerProTyrAlaAspGlnSer 599
Db 3689 CAGTACAACTGAGACCGCAACTCCGCTCGCTCA 3721

Search completed: May 26, 2003, 18:14:45
Job time : 494 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 18:03:19 ; Search time 90 Seconds
(without alignments)
2453.414 Million cells updates/sec

Title: US-09-978-385-2_COPY_19_738
Perfect score: 3869
Sequence: 1 STIEBQAKFLDKFNHREAD.....DNSLEFLGIQPTLPPNP 720

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+ p2n.model -DEV-xlh
-O-/cgn2.1/USPTO.spool/US09978385/rnat.21052003.154420.19853/app.query.fasta_1.903
-DB-Issued Patents.NA -OFMT-fastp -SUFFIX-rni -MINMATCH-0.1 -LOOPT-0
-LOOPT-0 -UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST-45 -DOCALLIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-20
-MODE-LOCAL -OUTFMT-pio -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USER-US09978385.ecgn.1.1.40.etrnat.21052003.154420.19853 -NCP-6 -ICPU-3
-NO_MMAP -LARGEOUTERY -NEG_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAEXT-0.5 -DELOP-6 -DELEXT-7

Database:

Issued Patents.NA:*
1: /cgn2.6/prodata/1/lna/5A.COMB.seq:*
2: /cgn2.6/prodata/1/lna/5B.COMB.seq:*
3: /cgn2.6/prodata/1/lna/6A.COMB.seq:*
4: /cgn2.6/prodata/1/lna/6B.COMB.seq:*
5: /cgn2.6/prodata/1/lna/PCTUS.COMB.seq:*
6: /cgn2.6/prodata/1/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the total score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3869	100.0	2415	4	US-08-989-299-3
2	3869	100.0	3396	4	US-08-989-299-1
3	3214	83.1	2350	4	US-09-280-116-40
4	1337	34.6	2478	1	US-08-481-626-1
5	1337	34.6	4020	4	US-09-050-159-130
6	1337	34.6	4024	4	US-09-162-484-18
7	1310	33.9	3942	4	US-09-162-484-19
8	721	18.6	2082	4	US-09-440-325A-2
9	387	10.0	467	1	US-08-664-596B-14
10	272	7.0	848	4	US-08-905-223-27
11	272	7.0	848	4	US-09-247-155-27
12	133.5	3.5	192	1	US-08-157-171-6

13	123	3.2	8700	2	US-08-392-625-16	Sequence 15, Appl
14	121	3.2	8700	2	US-08-466-961A-16	Sequence 16, Appl
15	121	3.1	789	4	US-09-280-116-114	Sequence 114, Appl
16	120.5	3.1	11384	4	US-08-961-527-45	Sequence 45, Appl
17	120	3.1	144	4	US-08-157-171-4	Sequence 4, Appl
18	120	3.1	11303	4	US-08-961-527-115	Sequence 115, Appl
19	119	3.1	6700	2	US-08-645-193B-118	Sequence 18, Appl
20	116	3.0	2052	3	US-08-630-916A-45	Sequence 18, Appl
21	116	3.0	2052	3	US-08-630-916A-45	Sequence 18, Appl
22	113.5	2.9	1052	4	US-09-453-702B-256	Sequence 108, App
23	113.5	2.9	1998	3	US-08-844-059-1	Sequence 1, Appl1
24	112	2.9	1839	3	US-09-431-202-1	Sequence 1, Appl1
25	112	2.9	1839	3	US-09-120-365-4	Sequence 4, Appl1
26	110	2.8	1839	4	US-09-515-039-4	Sequence 4, Appl1
27	109	2.8	13309	4	US-08-961-527-202	Sequence 202, App
28	108.5	2.8	9510	4	US-09-453-702B-256	Sequence 108, App
29	107.5	2.8	6506	4	US-09-453-702B-256	Sequence 256, App
30	107	2.8	3425	4	US-09-107-149-16	Sequence 16, Appl
31	105.5	2.7	37948	4	US-09-251-645-11	Sequence 11, Appl
32	104.5	2.7	3883	1	US-08-468-036-33	Sequence 33, Appl
33	104.5	2.7	3883	2	US-08-376-843-33	Sequence 33, Appl
34	104.5	2.7	3884	4	US-09-541-782-3	Sequence 3, Appl1
35	104.5	2.7	3884	4	US-09-723-820-3	Sequence 3, Appl1
36	104	2.7	3672	1	US-08-491-357-1	Sequence 1, Appl1
37	104	2.7	3672	3	US-08-968-633-1	Sequence 1, Appl1
38	104	2.7	3672	3	US-09-196-466-1	Sequence 1, Appl1
39	104	2.7	3672	5	PCT-US96-10823-1	Sequence 1, Appl1
40	101.5	2.6	1683	4	US-09-134-001C-316	Sequence 316, App
41	101.5	2.6	1944	2	US-08-844-056-1	Sequence 1, Appl1
42	101.5	2.6	3258	4	US-09-595-424-7	Sequence 7, Appl1
43	101.5	2.6	21706	4	US-08-961-527-36	Sequence 36, Appl
44	99.5	2.6	2793	4	US-08-953-040-1	Sequence 1, Appl1
45	99.5	2.6	10690	4	US-08-961-527-93	Sequence 93, Appl1

ALIGNMENTS

RESULT 1
US-08-989-299-3
; Sequence 3, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Robinson, Keith E.
; TITLE OF INVENTION: ANGIOGENESIS CONVERTING ENZYME HOMOLOG
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HONG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 317
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold E., Beth
; REGISTRATION NUMBER: 35,430
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-7000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2415 base pairs

Db 1975 AAAATACAGATGATCTTTTGGGAGGAGATGTCGAGTGCCTAATTTGAACCAACA 2034
 QY 115SerPheAsnPhenPheValThrAlaProLysAsnValSerAspIleIleProAlaThr 680
 Db 2035 ATCTCTTAAATTTCTTTGTCACATGCACTAAATAATGTCTGATATCTTCTAATACT 2094
 QY 681 GUAUAGLUysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 Db 2095 GAAGTGAAGAAAGGCATCAGATGATCCGAGCGCATCAATGATGCTTCCGCTGAAAT 2154
 QY 701 ASPAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGluProProAsnGlnProPro 720
 Db 2155 GACAAACAGCCTAGATTCTTGGGAGAAACACCAACTTGACCTCTTAACCAAGCCCT 2214

RESULT 2

US-08-989-299-1

Sequence 1, Application US/08989299

Patent No. 6194556

GENERAL INFORMATION:

APPLICANT: Acton, Susan L.

APPLICANT: Robinson, Keith E.

TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG

TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESSES:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,299

FILING DATE: 11-DEC-1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Arnold E., Beth

REGISTRATION NUMBER: 35,430

REFERENCE/DOCKET NUMBER: MIA-025.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-832-1000

TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3396 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 82..2496

US-08-989-299-1

Alignment Scores:

Pred. No.: 0 Length: 3396
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x US-08-989-299-1 (1-3396)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 Db 136 TCCACCATTTGAGGAAAGCCCAAGACATTTTGGACAGATTAAACACAGAGCCGAGAGAC 195

QY 21 LeuPheTyrglnSerSerLeuAlaSerTrpAsnTyraThrAsnIleThrGluGluAsn 40
 Db 196 CTGTTCTATCAAAAGTTCATCTGCTTCTTGAAATTAATACCAATATTAATGAAAGAGAT 255
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
 Db 256 GTCCAAACATGAATAATGATGCGGACAAATGGTGTGCTTTTAAAGCAACAGTCCACA 315
 QY 61 LeuAlaGlnMetTyrrProLeuGlnIuileGlnAsnLeuThrValLysLeuGlnLeuGln 80
 Db 316 CTGGCCAAATATATCCATACAAAGAAATTCAGAAATTCACAGTCAACCTCAGCTGACG 375
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 Db 376 GCTCTTCAGCAAAAGGGCTTCAGTGGCTGTCAGAGAACAGAACAGGCTTGACACCA 435
 QY 101 IleLeuAsnThrMetSerThrIleTyrrSerThrGlyLysValLysAsnProAspAsnPro 120
 Db 436 ATTCATAATACATGACATGACACATCTACATCTGGAAGAAAGTTTGTAACCCAGATATCCA 495
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrr 140
 Db 496 CAAGATGCTTTATTAATTGAAACAGGTTGAATGAATAATGCAACAGATTAGACTAC 555
 QY 141 AsnGluArgLeuThrAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
 Db 556 AATGAGAGGCTCTGGGCTTGGAAGCTGAGATCTGAGCTGGGACAGCAGCTGAGGCCA 615
 QY 161 LeuTyrglnGluTyrrValIleuLysAsnGluMetAlaArgAlaAsnHisTyrrGluAsp 180
 Db 616 TTAATATGAAGATATGATGCTTTGAAATAATGATGCGCAAGACCAATATATAGAGAC 675
 QY 181 TyrglyAspTyrrTrpArgLysAspTyrgluValAsnGlyValAspGlyTyrrAspTyrrSer 200
 Db 676 TATGGGATTAATGGAGAGAGACATGATGAAGTAATGGGGTATGAGGCTATACACAGC 735
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGlnIuileLysProLeuTyrrGlu 220
 Db 736 CGCGGCCAGTGTATTAAGATGTGAACATACCTTTGAAAGAAATTAACCATTAATATGAA 795
 QY 221 HisLeuHisAlaTyrrValArgAlaLysLeuMetAsnAlaTyrrProSerTyrrIleSerPro 240
 Db 796 CATCTTCATGCGCATGTGAGGGGCAAAAGTGAATGAATGCTTCCATATCATGTCGA 855
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrrGlyArgPheTrpThrAsnLeu 260
 Db 856 ATGGATGCCCTCCCTGCTCATTTGCTGGTGAATGTGGGGTATGGAATTTGGACAAATCTG 915
 QY 261 TyrrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 Db 916 TACTCTTGGACAGTTCCTTTGGACAGAAACCAAAACATGAGTTACTGATGCAATGGTG 975
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 Db 976 GACGAGGCCCTGGGATGACAGAGAAATTAATCAAGAGCCGAGAAAGTCTTTGATCTGTT 1035
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
 Db 1036 GGTCTTCCCTTAATATGACTCAAGGATTCCTGGGAAATTCATGCTAACGAGCCAGGAAT 1095
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 Db 1096 GTTCGAAACAGATGTGCAATCCACAGCTTGGGACCTGGGAAAGGCCATTCACAGATC 1155
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHis 360
 Db 1156 CTATATGCGCAAAAGGACAAATGAGACGACTCTGACAGCTCATCATGAGATGCGGCGAT 1215
 QY 361 IleGlnTyrrAspMetAlaTyrrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 Db 1216 ATCCAGATATATATGATATGCTGCAACCTTTCTGTAAAGAAATGAGCTAATATGAA 1275
 QY 381 GlyPheHisGluAlaValGlyIuileMetSerLeuSerAlaAlaThrProLysHisLeu 400

```

Db      1276 GGATTCAGTGAAGCTGGTGGGAATCATGTCACCTTGTGAGCCACACCTAACCATTTA 1335
Qy      401  LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAspPhe 420
Db      1336 AATCCATTTGGTCTCTGTCAACCCGATTTTCAAGAGACATGAAACGAATTAATTAATTC 1395
Qy      421  LeuLeuGlyGlnAlaLeuThrIleValGlyThrLeuProPheThrThrMetLeuGluLys 440
Db      1396 CTGCCTCAACAGACGCTCCACATGTTGTCGACCTGTCATTTACTTACATGTATGAGAG 1455
Qy      441  TrpArgTrpMetValPheLeuGlyGluIleProLysAspGlnTrpMetLysLysTrp 460
Db      1456 TGGAGGTGATGGTCTTTAAAGGGGAATTCCTCAAGACAGTGGATGAAAAAGTGGTGG 1515
Qy      461  GluMetLysArgGluIleValGlyValGluProValProHisAspGluThrThrCys 480
Db      1516 GAGATGAAGCAGAGATAGTGGGTGGTGGACCTGGCCCATGATGAAACATTAATCTGT 1575
Qy      481  AspProAlaSerLeuPheHisValSerAsnAspLysSerPheIleArgTrpThrArg 500
Db      1576 GACCCCGCATCTGCTTCATGATGTTCTAATGATTACATTCATTCATTCATTCATTCAT 1635
Qy      501  ThrLeuTrpGlnPheGlnPheGlnGluAlaLeuCysGlnAlaIleAspGluThrThrCys 520
Db      1636 ACCCTTACCATTCAGATTCATTCAGAACGACCTTTGTACAGCAGCTTAACATGAAAGCCCT 1695
Qy      521  LeuHisLysCysAspLysSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
Db      1696 CTGCACAAATGTGACATCTCAACCTACAGAACGCTGACAGAACGCTTCATTAATATCTG 1755
Qy      541  ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuLysValValGlyAlaLysAsn 560
Db      1756 AGCCTTGGAAATATGAAACCTGAGCCCTTACCATTTGGAATGTTGAGAGCAAAAC 1815
Qy      561  MetAsnValArgProLeuLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAspGln 580
Db      1816 ATGAATGTAAAGGCACGCTCACTACTTGTGAGCCCTTATTTACTTGCTGAAAGACAG 1875
Qy      581  AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTrpAlaAspGlnSerIle 600
Db      1876 AACAGAAATCTTTTGGGATGAGTACGACGAGCTGACATGACAGCAAAAGCATC 1935
Qy      601  LysValArgIleSerLeuLysSerAlaLeuGlyAspValAlaTrpValTrpAsnAspAsn 620
Db      1936 AAGGTGAGCATTAACCTTAATACGCTCTTGAGATTAACCATTAATGAAAGCAACAT 1995
Qy      621  GluMetTrpLeuPheArgSerSerValAlaTrpAlaMetArgGlnTrpPheLeuVal 640
Db      1996 GAATGTACTGTTCCGATCATCTGTGATATGCTATGAGCAGTACTTTTAAAGTA 2055
Qy      641  LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
Db      2056 AAAAATCAGATGATCTTTTGGGAGAGAGAGTGGAGTGGCTAATTTGAAACCAAGA 2115
Qy      661  IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgTrp 680
Db      2116 ATCCCTTAAATTTCTTTGTCACCTGACCACTAAATAATGCTGATATTCATCTTAAGACT 2175
Qy      681  GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
Db      2176 GAAGTTGAAAGGCCATCAGAGATGTCGCGAGCCGATCAATGATGCTTCCGCTGAAAT 2235
Qy      701  AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
Db      2236 GACACAGCCTTAAGTTTCTGGGAGATACACCAACACTTGGACCTCTTAACAGCCCTCT 2295

```

RESULT 3
 US-09-280-116-40/c
 ; Sequence 40. Application US/09280116A
 ; Patent No. 6331427
 ; GENERAL INFORMATION:
 ; APPLICANT: Robison, Kelch E.

```

; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280, 116A
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 40
; LENGTH: 2350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: angiotensin-converting enzyme
; US-09-280-116-40

Alignment Scores:
Pred. No.: 0 Length: 2350
Score: 3214.00 Matches: 706
Percent Similarity: 92.55% Conservative: 2
Best Local Similarity: 92.29% Mismatches: 12
Query Match: 83.07% Indels: 45
DB: Gaps: 2

US-09-978-385-2_COPY_19_738 (1-720) x US-09-280-116-40 (1-2350)
Qy      1  SerThrIleGluGlnAlaLysThr-PheLeuAspLysPheAsnHisGlu-AlaGlu 20
Db      2250 TCCACCATGAGAGAACAGCCCAAGACATTTTGGACAGATTAAACCAAGAACGCCAAG 2191
Qy      20  spleuPheTrpGlnSerSerLeuAlaSerTrpAsnTrpAsnThr-AsnIleThrGlu 39
Db      2190 ACCTGTTTATCAAAAGTACTGCTCTTGGAAATTAACCCCAATTAATTAATTAATTA 2131
Qy      40  AsnValGlnAsnMetAsnAsnAlaGlyAsp-LysTrpSer-AlaPheLeuGluGln 59
Db      2130 AATGTCCAAACATGATATATCTGGCGACAAATGCTGCTCTTTTAAAGACAGT 2071
Qy      59  erThrLeuAlaGlnMetTrpProLeuGlnGluIleGlnAsnLeuThrValLysLeuGln 79
Db      2070 CCACACTTCCCAAAATGATCCACTACAAAGAAATTCAGATCTCAGAGTTCATC 2011
Qy      79  euGlnAlaLeuGlnGlnGlnGlnSerValLeuSerGluAspLysSerLysArgLeu 99
Db      2010 TCGAGCTCTTCAACAAATGGGTCTCACTGCTCTCAAGAACAAAGCAACGCTTGA 1951
Qy      99  snThrIleLeuAsnThrMetSerThrIle-TrpSerThrGly-LysValCysAsnProAs 118
Db      1950 ACACAAATCTTAATTAATCAATGACACACATCATCACTACTGGAAGAAAGTTTAAACCCAGA 1891
Qy      118  pAsnProGlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerIle 138
Db      1890 TAAATCCCAAGAAATCTTAATTAATCAATGACACACATCATCACTACTGGAAGAAAGTTT 1831
Qy      138  uAspTrpAsnGluAspLeuTrpAlaTrpLysSerTrpArgSerGluValGlyGlnLe 158
Db      1830 ACACATCAATGAGAGCTCTGGGCTTGGGAAAGCTGGGAGATCTGGGAGCAAGCACT 1771
Qy      158  uArgProLeuTrpGluGluTrpValValLeuLysAsnGluMetAlaArgAlaAsnHisTrp 178
Db      1770 GAGGCCATTAATTAAGAGATAGTGTCTTGAATAATGAGATGCAAGCAAGCAATCATTA 1711
Qy      178  rGluAspTrpGlyAspTrpTrpArgGlyAspTrpGlu-ValAsnGlyVal-AspGlyTrp 197
Db      1710 TGAGGACTATGGGATTTATGGAGAGAGACTATGAAATGAGGCTTAAGATGAGCTAT 1651
Qy      198  AspTrp-SerArgGlyGlnLeuIleGluAspValGluHisThr-PheGlnGluIleLysP 217
Db      1650 GACTACATAGCCGCGCCGAGTGTGATTAAGATGTGCAACATACCTTTTGAAGATTAAC 1591
Qy      217  rGluTrpGluHisLeuHisAlaTrpValArgAlaLysLeuMetAsnAlaTrpProSer 237
Db      1590 CATATATGAACATCTTCATGCTATGTCAGAGGCAAAAGTTGATGAATGCCATCTCTCT 1531
Qy      237  yTrpLeuSerProIleGlyCysLeuProAlaHisLeu-LeuGlyAspMetTrpGly-Argph 256

```

Db 1530 ATTCAGTCCATTTGGATGGCTCCCTGCTCATTTCCGTTGGATATGGGGTTAGATT 1471
 QY 256 eTRP-ThrasnLeuTySerLeuThrVal-ProPheGlyGln-LysProAsnIle-AspY 275
 Db 1470 TTGGGCAAAATCTGTACTCTTTGACAGTGTCCCTTTGGACAGAAACAAACATAGATG 1411
 QY 275 alThraspala-MetValAspGlnAlaTTPAspala-GlnArgIlePheGlyGlnAla-G 294
 Db 1410 TTACTGATCCCAATGGGACAGCCCTGGATGACACAGAAATATTCAGAGAGCCCG 1351
 QY 294 IulysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheTrp-GluAsnSer 313
 Db 1350 AGAAGTCTTGTGATCTGTGGCTCTCTTAATATGACTCAAGATTCCTGGGAAAAATTC 1291
 QY 314 MetLeuThr-AspProGlyAsnValGlnLysAlaValCysHisProThrAlaTTPAsp 333
 Db 1290 ATGTGAACGGGACCCAGAGAAATGTTCAGAAAGCAGTCTGCCATCCACAGCTTGGAGCT 1231
 QY 333 u-GlyLysGlyAsp-PheArgIle-LeuMetCysThrLysValThrMetAspPheLe 352
 Db 1230 GGGGGAGGCGGAGCTTTCAGATCTTAAATGTGCACAAAGGTGACATGGAGCAGCTTCT 1171
 QY 352 uThrAlaHisHisGlyMetGlyHisIleGlnTyraPheMetAlaTyrAlaIleGlnPro 372
 Db 1170 GACAGCTCATCATGAGTGGGATGCGATATGATGATGATGATGATGATGATGATGATG 1111
 QY 372 eLeuLeuArgAsnGlyAlaAsnGlnGlyPheHisGlnAlaValGlyGlnIleMetSer 392
 Db 1110 TCTGGTAAGAAATGGAGCTAATGAGATTCATGAAAGCTTTGGGAAATCATGTCAC 1051
 QY 392 uSerAlaIleThrPro-LysHisLeuLysSerIleGlyLeuLeu--SerProAspPheG 411
 Db 1050 TTTCGCGACGACACCTTAAACATTTAAATCCCATGCTGTTTGTTCACCCAAATTTTC 991
 QY 411 InGluAspAsnGlnuThrGluIleAsnPheLeuLysGln-AlaLeuThrIleValGly 430
 Db 990 AAGAGACATGAAACGAAATTAATCTCTGCTCAAAACAAACACACAGATGTTGGG 911
 QY 431 Thr-LeuProPheThrTyMetLeuGlnLysTrpArgTrpMetValPheLysGlyGln 450
 Db 930 ACTGCTGCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 871
 QY 450 eProLysAspGlnTrpMetLys-LysTrpTrpGluMetLysArgGluIleValGlyVal 470
 Db 870 TCCCAAGACACAGTGGATGAAACAAAGTGTGGAGATGAGACAGATGATGATGATG 811
 QY 470 al---GluProVal-ProHisAspGlnuThrTyraPheAspProAlaSerLeuPheHisVal 488
 Db 810 TTGGGAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 751
 QY 489 SerAsnAspTySer-PheIleArgTyTrpArgTrpLeuThrGlnPhe-GlnPhe 507
 Db 750 TCTAATGATTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 651
 QY 508 GlnGluAlaLeuLys-GlnAlaAlaLysHisGlnGlyProLeuHisLysCysAspIle 527
 Db 650 CAAGAACACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 631
 QY 527 r-AsnSerThrGlnAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGlu 547
 Db 630 GTAACTCTACAGAGCTGGACAGAAACTGTCAATATGCTGAGGCTTGGAAAAATCAGAAC 571
 QY 547 rOTrPheLeuAlaLeuGlnValAsnValAlaGlyAlaLys-AsnMetAsnValArgProLeu 566
 Db 570 CCGGAGACCTAGACATTGGAATATGTTAGAGCAAAAGAAACATGATGATGATGATG 511
 QY 567 LeuAsn-TyrPheGlu-ProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 586
 Db 510 CTCAACTACTTGTGAGCCGCTTATTTACCTGCTGAAAGACAGAACAAATTTCTTTG 451
 QY 586 alGlyTrpSerThrAspTrpSerProTyraIleAspGln-SerIleLysValAlaGlyLe 605

Db 450 TGGATGAGTACGACGAGTGGATCCATATGACAGACCAAAAGCATCAAAATGAGATTAAG 391
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyraGluTrpAsnAspAsnGluMetTyraLeu 625
 Db 390 CCTAAATACGCTTGTGGAGATTAAGCATATGATGAAACGACATGAAATGTACTCTTT 331
 QY 625 eArgSerValAlaIleTyrAlaMetArgGlnTyraPheLeuLysValLysAsnGln-Met 645
 Db 330 CCGATCATCTGTTGATATGATATGAGGAGCAGTCTTTTAAAGTAAATAATCAGACATGA 271
 QY 645 leuPheGlyGlnGlu-AspValArgValAlaAsnLeuLysProArgIleSerPheAsn 664
 Db 270 TTCTTTTGGGAGAGAGATGTGCGAGGCTTAATTTGAAACCAAGATCTCTTTAT 211
 QY 665 PhePhe-ValThrAlaProLysAsnValSerAspIleLeuProArgTrpGlnValGly 684
 Db 210 TTCTTTGGTACTGCTCCCTTAAATATGTCTGATATCATCTCTAATCTGAAAGTTGAAA 151
 QY 684 salalIeArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSer 704
 Db 150 GGCATCAGAGATGCTCCGAGCCGATCATGATGCTTCCGCTGAATGACACAGCT 91
 QY 704 uGluPheLeuGlyIleGlnProThrLeuGlnProProAsnGlnProPro 720
 Db 90 AGAGTCTTGGGATACAGCCAAACCTTGAGCTCTTAACAGCCCT 42
 RESULT 4
 US-08-481-626-1
 : Sequence 1, Application US/08481626
 : Patent No. 5801040
 : GENERAL INFORMATION:
 : APPLICANT: Soudrier, Florent
 : APPLICANT: Albenc-Gelas, Francois
 : APPLICANT: Hubert, Christine
 : APPLICANT: Corvol, Pierre
 : TITLE OF INVENTION: Nucleic Acid Coding for the Human
 : TITLE OF INVENTION: Testicular Angiotensin Converting Enzyme (ACE) and its
 : TITLE OF INVENTION: Uses, Especially for the In Vitro Screening for this
 : NUMBER OF SEQUENCES: 3
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Flanagan, Henderson, Farbow, Garrett &
 : ADDRESSEE: Dunner
 : STREET: 1300 I Street, N.W., Suite 700
 : CITY: Washington
 : STATE: D.C.
 : ZIP: 20005-3315
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/481,626
 : FILING DATE: 07-JUN-1995
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/656,183
 : FILING DATE: 04-MAR-1991
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: FR 89-09062
 : FILING DATE: 05-JUL-1989
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Meyers, Kenneth J.
 : REGISTRATION NUMBER: 25,146
 : REFERENCE/DOCKET NUMBER: 04958-0006-02000
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 202-408-4000
 : TELEFAX: 202-408-4400
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 2478 base pairs

||||| :||||| :|||||
Db 3629 CCGCTGCTGAGTGGCTCCGACGAGAGACGCTGATGGAGAGAGCTGGCTGGCCG 3688
QY 589 SerThrAspTyrPheProTyrAlaAspGlnSer 599
Db 3689 CAGTACACTGAGCCGCACTCCGCTGCTCA 3721
RESULT 6
US-09-162-484-18
Sequence 18, Application US/09162484
Patent No. 6248724
GENERAL INFORMATION:
APPLICANT: Phillips, M. Ian
APPLICANT: Mohuczy, Dagmara
TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS TARGETED TO
TITLE OF INVENTION: ANGIOSENSIN CONVERTING ENZYME MRNA AND METHODS OF USE
FILE REFERENCE: UFLA:087/DEL087P 484
CURRENT APPLICATION NUMBER: US/09/162,484
CURRENT FILING DATE: 1998-09-25
EARLIER APPLICATION NUMBER: 60/059,661
EARLIER FILING DATE: 1997-09-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 18
LENGTH: 4024
TYPE: DNA
ORGANISM: Homo sapiens
US-09-162-484-18
Alignment Scores:
Pred. No.: 1,256-149 Length: 4024
Score: 1337.00 Matches: 255
Percent Similarity: 61.05% Conservative: 118
Best Local Similarity: 41.73% Mismatches: 204
Query Match: 34.56% Indels: 34
Dbs: 4 Gaps: 9
US-09-978-385-2_COPY_19_738 (1-720) x US-09-162-484-18 (1-4024)
QY 2 ThrileugluginAlaLysThrPheLeuAspPheAsnHisGluAlaGluAspLeu 21
Db 1952 ACTGATGAGGCTGAGGAGCCAGCAAGTTGTGAGAGAAATGACACCGACATCCAGGTGTG 2011
QY 22 PheTyrGlnSerSerLeuAlaSerTyrPheAsnThrAsnThrAsnThrGlnGlu----- 39
Db 2012 TGGAGCAGATGCGCGAGCCAACTGACATCAACCAATCCATCCACAGACAGCAGC 2071
QY 40 -----AsnValGlnAsnMetAsnAlaGlyAspTyrSerAla 53
Db 2072 AAGATTCGTGCTGACAGAAACATGCAATACCAACACAC----- 2113
QY 54 PheLeuLysGlnGlnSerThrLeuAlaGlnMetTyrProLeuGlnGlnLeuGlnAsnLeu 73
Db 2114 -----CTGAGTACGCGACCCAGCCAGGAAAGTTGATGTAACCACTTCAGAACACC 2167
QY 74 ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGlnAsp 93
Db 2168 ACTATCAAGCGGATCAAGAAAGGTTCAGACCTAGAACCGGACGCTCCGCCAG 2227
QY 94 LysSerLysArgLeuAsnThrThrLeuAsnThrMetSerThrThrLysSerThrGlyLys 113
Db 2228 GAGCTGGAGAGTACACAAAGATCTCTGTGATATGAAACACACCTACAGCTGGCCACT 2287
QY 114 ValCysAsnProAsnProGlnGlnCysLeuLeuLeuGlnProGlyLeuAsnGlnLeu 133
Db 2288 GGTGTCACACCGAAATGCG-----AGCTGCGTGGACGCTGAGCCAGATCTGACGAATGTG 2341
QY 134 MetAlaAsnSerLeuAspTyrAsnGlnGlnLeuThrAlaTyrGlnSerTyrAspSerGln 153
Db 2342 ATGGCCACATCCCGAAATATGAAGACCTGTATGGCATGGAGGGCTGGCGAGACAG 2401
QY 154 ValGlyLysGlnLeuArgProLeuTyrGlnGlnTyrValValLeuLysAsnGlnMetAla 173
||||| :||||| :|||||

Db 2402 GCGGAGAGAGCCATCTCCAGTTTACCCGAAATACCTGGAATCATCAACGAGCTGCC 2461
QY 174 ArgAlaAsnHisTyrGlnAspTyrGlnAspTyrTyrPheGlyAspTyrGlnValAsnGly 193
Db 2462 GGGCTCAATGCTATGTAGTGCAGGGAGCTGGAGGTCTATGTACAGACACCATCC 2521
QY 194 ValAspLysTyrAspTyrSerArgGlyGlnLeuLeuGlnAspValGlnHisThrPheGln 213
Db 2522 CTGGAG-----CAAGACCTGGAGGGCTCTCCAG 2551
QY 214 GlnLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAla 233
Db 2552 GAGCTGAGCCACTTCACTCACTGACATGCTGAGTGGCGGGCCCTGACCGCTCAC 2611
QY 234 Tyr-----ProSerTyrLeuSerProLeuGlyCysLeuProAlaHisLeuLeuGlyAspMet 252
Db 2612 TACGGGGCCAGACATCAACCTGGAGGGCCCATCTCTGCTACCTGCTGGGGAATG 2671
QY 253 TrpGlyArgPheThrPheAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsn 272
Db 2672 TGGCGGAGACTGTGTCAAATCTATGACTGTGTGCTCCCTCCCTGACCGCCCTCG 2731
QY 273 IleAspValThrAspAlaMetValAspGlnAlaTyrAspAlaGlnArgLeuPheLysGln 292
Db 2732 ATGGACACCCACAGAGCTATGCTAAAGCAGGGCTGAGCCGCCAGAGAGATGTTAAAGAG 2791
QY 293 AlaGlnLysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheThrGlnAsn 312
Db 2792 GCTGATGATTTCTTCACTCCCTGGGGCTGCTCCGCTCCCTGATGTTGGAACAG 2851
QY 313 SerMetLeuThrAspProGlnAsnValGlnLysAlaValCysHisProThrAlaTyrAsp 332
Db 2852 TCGATGTGGAGAAAGCCAAACCGAGGGCGGAGAGTGTGTCACCCCTCGCGCTGGAGC 2911
QY 333 LeuGlyLysGly-----AspPheArgIleLeuMetCysThrLysValThrMetAspPhe 351
Db 2912 TTCTACACGCGCAGAGACTTCGCGATCAAGCACTGACACCACTGACCTGGAGACCTG 2971
QY 352 LeuThrAlaHisHisGlnMetGlnHisIleGlnTyrAspMetAlaTyrAlaGlnPro 371
Db 2972 GTGGTGGCCACACGAAATGGGCGCACATTCGATGAAATGATGGCCCTTACAAAGACTTACT 3031
QY 372 PheLeuLeuArgAsnGlyAlaAsnGlnGlyPheHisGlnAlaValGlyLysLeuMetSer 391
Db 3032 GTGGCCCTGGAGGAGGTGTCACACCCGGCTTCATAGAGGCAATTTGGAGAGCTGAGCC 3091
QY 392 LeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGln 411
Db 3092 CTCTCAGTGTCTACGCGCAAGCAGCTGACAGCTGACACTGCTGTGACAGTGGAGGTGCG 3151
QY 412 GlnAspAsnGlnThrGlnLysAsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThr 431
Db 3152 AGCGAC-----GAGCATGACATCAACTTTCGATGAAATGATGGCCCTTACAAAGACTTCT 3208
QY 432 LeuProPheThrTyrMetLeuGlnLysTyrPheTyrMetValPheLysGlyLysPro 451
Db 3209 ATCCCTCTGAGCTACCTCGTGCATCAGTGGCGGTGAGGATTTATGAAAGCATCACCC 3268
QY 452 LysAspGlnThrMetLysTyrTyrGlnMetLysArgGlnLysValValValGln 471
Db 3269 AAGGAGAACTATTAACAGAGAGGTGTGAGGCTGAGGCTGAAGTACAGAGGCGCTGCCCC 3328
QY 472 ProValProHisAspGlnThrTyrCysAspProAlaSerLeuPheHisValSerAsnAsp 491
Db 3329 CCAATGCCCGAGACTCAAGTGTGACTTTCACCCCGAGGCGAAAGTTTCCACATTCCTCTTAC 3388
QY 492 TyrSerPheIleArgTyrTyrThrArgTyrThrLeuTyrGlnPheGlnGlnGlnAlaLeu 511
Db 3389 GTGCTTACATCAAGGATCTTGTACACTTATCATCTCAATCTCAAGTTCACAGAGGAGCTG 3448
QY 512 CysGlnAlaAlaLysHisGlnGlnGlnProLeuHisLysCysAspIleSerAsnSerThrGln 531
Db 3449 TGGCAGGACACTGTGCCACAGGCGCCCTGCAACATGTGACATCTACCAAGTCCAAAGAG 3508

QY 532 Alaglylnlyseupheasmetleuargleuglylysergltuprottrphtlreua 551
 DB 3509 GCGGGGAGCGCGTGGGACCGCCATACCTGGCTCAGTAGGCGCGCGGAGG 3568
 QY 552 LeuglnsvalValglalValysAsmetAsnValarProleuAsnlyrphleglu 571
 DB 3569 ATGCAGCTGATACAGCGCCAGCCACATACCGCCCTCGCCATGTTGACCTACTTCAG 3628
 QY 572 ProleuphettrphtleuLyAspGlnAsnlyS-----AsnSerPheValglYtrp--- 588
 DB 3629 CCGCTGCTGAGCTGCTCCGACGAGAACGAGCTGATGGGAGAGAGCTGGCTGGCG 3688
 QY 589 SertrasprrpserprotyrAlaAspGlnSer 599
 DB 3689 CAGTACAACTGAGAGCGGAACTCGCTGCTCA 3721

RESULT 7
 US-09-162-484-19
 : Sequence 19, Application US/09162484
 : Patent No. 6248724
 : GENERAL INFORMATION:
 : APPLICANT: Phillips, M. Ian
 : APPLICANT: Mohuczy, Dagnara
 : TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS TARGETED TO
 : TITLE OF INVENTION: ANGIOGENESIS CONVERTING ENZYME mRNA AND METHODS OF USE
 : FILE REFERENCE: DPLA:087/DPLA087P
 : CURRENT APPLICATION NUMBER: US/09/162,484
 : EARLIER FILING DATE: 1998-09-25
 : EARLIER APPLICATION NUMBER: 60/059,661
 : NUMBER OF SEQ. ID NOS.: 19
 : SOFTWARE: Patent In Ver. 2.0
 : SEQ ID NO 19
 : LENGTH: 3942
 : TYPE: DNA
 : ORGANISM: Rattus norvegicus
 : US-09-162-484-19

Alignment Scores:
 Pred. No.: 2,1e-146 Length: 3942
 Score: 1310.00 Matches: 251
 Percent Similarity: 61.20% Conservative: 115
 Best Local Similarity: 41.97% Mismatches: 214
 Query Match: 33.86% Indels: 18
 Gaps: 7

US-09-978-385-2_COPY_19_738 (1-720) x US-09-162-484-19 (1-3942)

QY 2 ThrllleuglulglulalalysThrphleuAspLyPheasnhlsGluAspLeu 21
 DB 1948 ACTGATGAAAGCCAGGCTAACAGGCTTCGTGAGAGTATGACCGGACCCAGGAGTGTG 2007
 QY 22 PheyrGlnSerSerleuAlaSerTrpAsnTrpAsnThrAsnlllethrgluGlnuVal 41
 DB 2008 TGAAGCAATATACGAGCGGACGACATTAATACACCAACTTTCATATAGAGGCGAC 2067
 QY 42 GlnAsnMetAsnAsnAlaGlyAspLySTrpSerAlaPheleuLySgluGlnSerThrleu 61
 DB 2068 AATACCTCTGCTTCAGAAAAAAGAAAGAGTCCCAACCATATCCCTTGAATATGACCTGG 2127
 QY 62 AlaGlnMetLyPheleuGlnGluValleuAsnLeuThrVallyrleuGlnleuGlnAla 81
 DB 2128 GCCAAGACATTTACGTGACCAACTTCAGAACTTACATCAAGCGGATCANTAAAGAG 2187
 QY 82 LeuGlnGlnAsnGlySerValleuSerGluAspLySerLySargLeuAsnThrle 101
 DB 2188 GTTCAGAACGTGAGCGGAGTGTCCCTCCCAAGAGATTAGAAAGATACCAACCAATC 2247
 QY 102 LeuAsnTrpMetSerThrlyrSerThrGlyLyValCysAsnProAspAsnProGln 121
 DB 2248 CTGCTAGACATGTGAGAGCACTTACACTGTAGCCAAATGTTTGTACACAAATAGGC----- 2301

QY 122 GluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerleuAspLyAsn 141
 DB 2302 ACTGTCTCTACGTGAGCGCTGATCTGACAAATATATAGGACACCTCCCGGAATATGCA 2361
 QY 142 GluArgLeuThrAlaTrpGlnSerTrpAspSerGluValGlyGlnleuArgProleu 161
 DB 2362 GAATTCCTTGTGGGTGTGGAGAGCTGGGAGCAAGAGTGGGAGAGCCATCTTCCCTTT 2421
 QY 162 TyrGlnGluTyrValValleuLySAsnGlnMetAlaArgAlaAsnHlsTyrGlyAspLy 181
 DB 2422 TTCCCAAGTACGTGGAGCTTCTCCAAAGATCGGCAAGCTACAGGCTACTCTGATCA 2481
 QY 182 GlyAspLyTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspLySerArg 201
 DB 2482 GGGGATTCCTGGAGATCTCATATGATGATGATGATGATGATGATGATGATGATGATG 2523
 QY 202 GlyGlnleuIleGluAspValGlnHlsThrPheGlnGluIleLyPheleuTyrGlnHls 221
 DB 2524 -----CAAGACCTGGAAAACTATACAGAGCTGACCGCTTACCTGAC 2571
 QY 222 LeuHlsAlaTyrValArgAlaLySLeuMetAlaAlaTyrProSer---TyrIleSerPro 240
 DB 2572 CTGCATGCTTATGTGGCGCGCTCCGACCGCATTTATGGGTGAGTACATCAACCTG 2631
 QY 241 IleGlyCysLeuProAlaHlsleuLeuGlyAspMetTrpGlyArgPheTrpThrAsnleu 260
 DB 2632 GATGGTCCATCTGCTCCGACCTGCTAGGGAACATGGGACACAGACTGTGTCACAACTC 2691
 QY 261 TyrSerleuThrValProPheGlyGlnLySProAlaIleAspValThrAspAlaMetVal 280
 DB 2692 TATGACTGTGGGACACCTTCCCTCCGCGCCCACTATAGAGGACCGAGGAGCATGTA 2751
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLySgluAlaGluLySPhleuValSerVal 300
 DB 2752 AAGCAGGAGATGACACCCAGAGAGATATTAAAGAGTGAACATTTTTCCTCCCTG 2811
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetleuThrAspProGlyAsn 320
 DB 2812 GGGCTGTACCTGCTCCGCTGAGTTCGAAACATGTTGAGAACCAACGAT 2871
 QY 321 ValGlnLySAlaValCysHlsProThrAlaTrpAspLeuGlySglY---AspPheArg 339
 DB 2872 GGGAGGAGAGTGTGCTCCATGCTCCAGCTGGGAGCTTTCACAAAGGAGAGACTCGG 2931
 QY 340 IleLeuMetCysThrLySValThrMetAspPheleuThrAlaHlsHlsGlnMetGly 359
 DB 2932 ATCAAGCAATGTACCTGTGTGAACATGGAGGATATGTGATAGGCCACAGAAATGGGC 2991
 QY 360 HlsIleGlnTyrAspMetAlaTyrAlaAlaGlnProPheleuLeuArgAsnGlyAlaAsn 379
 DB 2992 CACATCCAGTATTTATGATGACAGTACAAAGACTGTGCTGACCTTCGGGAGGCGCAAC 3051
 QY 380 GlnGlyPheHlsGlnAlaValGlyIleMetSerleuSerAlaAlaThrProLySAs 399
 DB 3052 CCGGCTTTTCAGAGGCTATTTGAGATGTTTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 3111
 QY 400 LeuLySserIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsn 419
 DB 3112 CTACACAGCTTCACACCTGCTGACAGCTGAG---GGCAGTGGCTACGACATGATCATC 3168
 QY 420 PheleuLeuLySAlaAlaLeuThrIleValGlyThrleuProPheThrTyrMetleuGlu 439
 DB 3169 TTCTAATGAAGATGGCTTTCAGAAAGTCCCTTCAATCCCTTCAAGCTACCTCATGAC 3228
 QY 440 LySTrpArgTrpMetValPheLySgluIlePheLySAspGlnTrpMetLyLyTrp 459
 DB 3229 CAGTGGCGCTGGAGGCTTTTACGAGACATCAACAGAGAACTACACACAGAGAGG 3288
 QY 460 TrpGlnMetLySArgGluIleValGlyValValGluProValProHlsAspGluThrTyr 479
 DB 3289 TGAAGTCTACAGACTAAGTACAGAGGCTGTGCTCCCTCAGAGGCTAGATCCCAAGGAGAC 3348
 QY 480 CysAspProAlaSerleuPheHlsValSerAsnAspLyTrpSerPheIleArgTyrTyrThr 499


```

Db      1420 YTNATGAAVYNTTYMGNTGAGAGTNTTYGAYGACNACNATHGGAAGATVHTTAYAA 1479
      457 LysLysTTrpTrpLuu---MetLysArgGluIleValGluValGluProValProHis 475
      1480 CARGARTGGTGAAYTNNMGNTNNAATATACAGACGTTNTGYCCNCCNATHCCNCAV 1539
      476 AspGluThrTyrCysAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIle 495
      1540 WMSGAGARAGATYTTGATCCNGGNCNCAATTTTCATTWMSNCCNGGNCNCCNTNAYAH 1599
      496 ---ArgTyrTTrpThrArgThrLeuTyrGlnPheGlnPheGlnGluAlaLeuGluAla 514
      1600 MGNMGNTAYTYYTNNMNTNNTNNTNCAATTCATTCATTCATTCATTCATTCATTCATTC 1659
      515 ALALysHisGluGluProLeuHisLysCysAspIleSerAsnSerThrGluAlaGluGln 544
      1660 WMSGNCATATGCGNCCNTNCAATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1719
      535 LysLeuPheAsnMetLeuArgLeuGlyLysSerGluProTyrPheLeuAlaLeuGluAsn 554
      1720 -----YTNTNCGNTNNAATYNTNGGWSNMAARCCNTGCCNGARGTNTNNAARATG 1773
      555 ValValGluAlaLysAsnMetAsnValArgProLeuAsnTyrPheGluProLeuPhe 574
      1774 YTNACNGCNGARNSNGARNTNMAACAAATGNTTAYAGACNTATTTAARCCNTNNTN 1833
      575 ThrTTrp 576
      1834 ACNTGG 1839

```

RESULT 9

```

US-08-664-596B-14
; Sequence 14, Application US/08664596B
; Patent No. 5807703
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallee, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; APPLICANT: Spaulding, Vikki
; APPLICANT: Bowman, Michael
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/664,596B
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-664-596B-14

```

```

Alignment Scores:
Pred. No.: 3,77e-37 Length: 467
Score: 387.00 Matches: 77
Percent Similarity: 98.72% Conservative: 0
Best Local Similarity: 98.72% Mismatches: 0
Query Match: 10.00% Indels: 1
DB: 1 Gaps: 0

```

US-09-978-385-2_COPY_19_738 (1-720) x US-08-664-596B-14 (1-467)

```

QY      1 SerThrIleGluGluGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
      236 TCCACCATTTGAGAACAGGACCAAGCATTTTGGACAGTTTAAACAGCAAGCCGAC 295
      21 LeuPheTyrGlnSerSerLeuAla-SerTyrAsnTyrAsnThrAsnIleThrGluGluAs 40
      296 CTGTTCTATCAAAAGTTCACTTCCTTGGATTTATAACACCATATTACTGAAGACAA 355
      40 nValGlnAsnMetAsnAlaGlyAspLysTyrSerAlaPheLeuLysGluGlnSerTh 60
      356 TGTCCAAACACGATTAATGCTGGGACAAATGGCTGCTTTTAAAGAACACTCCAC 415
      60 rLeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeu 77
      416 ACTTGCCCAATGATATCTACCTACAGAAATTCAGATTCACAGTCACAGCTT 467

```

RESULT 10

```

US-08-905-223-27
; Sequence 27, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 848 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:

```

FILING DATE: 24-JAN-1994


```

Db 2952 TTAGATCTTTAAAGCCGTCACACTTTGGAAAGCTT---ACGGAAATTTCAATATA 3008
Qy 192 AsnGlyValAspGlyTyrAspTyrSerArgGlyLeuLeuIleGluAspValGluHisThr 211
Db 3009 AAG-----AAAATAATCAATTAACAAAAGAAATGTCAT--- 3044
Qy 212 PheGluGluIleLysProLeuTyrGluHisLeuHisAlaTyrValAlaGlyLysLeuMet 231
Db 3045 -----CATTAACAATAATTAACGAATGAAGAAATGGTTTA 3077
Qy 232 AsnAlaTyrProSerTyrIleSerProIleGlyLysLeuProAlaHisLeuLeuGlyAsp 251
Db 3078 GAAATTAACCAATTAATGAAGGTCCTTACTACAGAAATGAATTAATTTTGAATTAAT 3137
Qy 252 MetTyrGlyArgPheThrPheAsnLeuTyrSerLeuThrValProPheGlyGluLysPro 271
Db 3138 -----AATAGAAATTAATTAATCTGTTAAATTAATTAATTAATTAATTAAT 3182
Qy 272 AsnIleAspValThrAspAlaMetValAspGluAlaThrAspAlaGluHisGlyLys 291
Db 3183 GATATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3242
Qy 292 Glu-----AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGluGly 308
Db 3243 GAAATTAACATGATTCAGAGATGCTATCTATCTAATCAATGTTAATTAATTAATTAAT 3302
Qy 309 -----PheThrGluAsnSerMetLeuThrAspPro 318
Db 3303 TCTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3362
Qy 319 GlyAsn-----ValGluLysAla 324
Db 3363 ATAACTGAGAGAGCATTGACTCATTAACCTTTTGTCCAGAAATTAATTAATTAATTAAT 3422
Qy 325 ValCysHisProThrAlaThrAspLeuGlyLysGlyAspPheArgIleLeuMetCysThr 344
Db 3423 ATTTTAAACACGACTTCTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3473
Qy 345 LysValThrMetAspAspPheLeuThrAlaHisIleGluMetGlyHisIleGlnTyrAsp 364
Db 3474 GAAATTAATGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3530
Qy 365 MetAlaTyrAla----- 368
Db 3531 GTAATTAATGCTTTTGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3590
Qy 369 -----AlaGlnProPheLeuLeuArgAsnGlyAla 378
Db 3591 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3650
Qy 379 AsnGluGlyPheHisGluAlaValGlyGluIleMetSer----- 391
Db 3651 AATGAATCTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3710
Qy 392 -----LeuSerAlaAlaThrProLys-----HisLeuLys 401
Db 3711 TTTAAAGAACATATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3770
Qy 402 -----SerIleGlyLeuLeuSerProAspPheGlnGluLysAsnGluThrGluIle 418
Db 3771 GATTGTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3815
Qy 419 AsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeu 438
Db 3816 AATTTTATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3875
Qy 439 GluLysTyrArgTyrMetValPheLysGlyGluIleProLysAspGlnTyrMetLysLys 458
Db 3876 AATTAATTAATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3923
Qy 459 TrpTrpGluMetLysArgGluIleValGlyValValGluProValProHisAspGluThr 478

```

```

Db 3924 -----TTAAGATTAATTAAGACA-----GATGAAGAT 3950
Qy 479 TyrCysAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyr 498
Db 3951 TAT-----TCTCAAAATTAATTTCTTCAATTAATTAATTAATTAATTAATTAATTAAT 3983
Qy 499 -----ThrArgThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCys 512
Db 3984 AAGATTAATTTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4043
Qy 513 GlnAlaAlaLysHisGluGlyProLeuHisLysCysAspIleSerAsnSerThrGluAla 532
Db 4044 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4103
Qy 533 GlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProThrPheThrLeuAlaLeu 552
Db 4104 GATAGTCTATTA-----TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4154
Qy 553 GluAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGluPro 572
Db 4155 GAAATTAATCGTGTCT-----ATATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 4202
Qy 573 LeuPheThrTyrLeuLysAspGlnAsnLysAsnSerPheValGlyTyrSerThrAsp 592
Db 4202 ----- 4202
Qy 593 SerProTyrAlaAspIleSerIleLysValArgIleSerLeuLysSerAlaLeuGlyAsp 612
Db 4203 -----ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4247
Qy 613 LysAlaTyrGluTyrPheAsnAspAsnGluMetTyrLeuPheArgSerSerValAlaTyrAla 632
Db 4248 GATTTATATCGTAGTATGAC----- 4268
Qy 633 MetArgGlnTyrPheLeuLysValLysAsnGlnMetIleLeuPheGlyGluLysAspVal 652
Db 4269 ATTAAGCAATAT----- 4280
Qy 653 ArgValAlaAsnLeuLysProArgIleSerPheAsnPhePheValThrAlaProLysAsn 672
Db 4281 -----AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4316
Qy 673 ValSerAspIleLeuPro----- 678
Db 4317 GACTATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4376
Qy 679 -----ArgThrGluValGluLysAlaIleArgMetSerArgSerArg 692
Db 4377 AGTATTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4436
Qy 693 IleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLeuGlyTyrIleGlnPro 711
Db 4437 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4490

```

RESULT 14
 US-08-466-961A-16
 Sequence 16, Application US/08466961A
 Patent No. 5843709
 GENERAL INFORMATION:
 APPLICANT: Entian, Karl-Dieter
 APPLICANT: Gtz, Friedrich
 APPLICANT: Schnell, No. 5843709bert
 APPLICANT: Augustin, Johannes
 APPLICANT: Engelke, Gerhart
 APPLICANT: Rosenstein, Ralf
 APPLICANT: Kaletta, Cortina
 APPLICANT: Klein, Cora
 APPLICANT: Mieland, Bernd
 APPLICANT: Kupke, Thomas
 APPLICANT: Jung, G nther
 APPLICANT: Kellner, Roland
 TITLE OF INVENTION: Biosynthetic Process for the Preparation of
 TITLE OF INVENTION: Chemical Compounds

NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 STREET: 1100 New York Avenue, NW
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/466,961A
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/392,625
 FILING DATE: 22-FEB-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/876,791
 FILING DATE: 30-APR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/784,234
 FILING DATE: 31-OCT-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Esmond, Robert W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 0652.0980004
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8700 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-466-961A-16
 Alignment Scores:
 Pred. No.: 0.00266 Length: 8700
 Score: 123.00 Matches: 149
 Percent Similarity: 32.36% Conservative: 116
 Best Local Similarity: 18.19% Mismatches: 250
 Query Match: 3.18% Indels: 304
 DB: 2 Gaps: 39
 US-09-978-385-2_COPY_19_738 (1-720) * US-08-466-961A-16 (1-8700)
 QY 2 ThrIleGluGlnAlaIstYrPhleuAspLysPheAsnHisGlu----- 17
 Db 2619 ACATATTAGAAATATACGAAATTTTATGATTAATATGATTAATGACACACAGTAAT 2678
 QY 18 -----AlaGluAspLeu-----PheTyrGlnSerSerLeuAlaSerTyrPasn 31
 Db 2679 TTAAGCAATTCCTCCAGATATTAATGATTTGGCTATCCCAAAAAGACGTTATAGT 2738
 QY 32 TYrAsnThrAsnIleThrGluGlnAsnValGlnAsnMetAsnAlaGlyAspLysTyr 51
 Db 2739 TTTTCTTAATACATTT----- 2753
 QY 52 SerAlaPheLeuLysGlnSerThrLeuAlaGlnMetLysProLeuGlnGluIleGln 71
 Db 2754 -----GCATTTTAAAGAAAAG----- 2771
 QY 72 AsnLeuThrValLysLeuGlnIleGlnAlaLeuGlnIleAsnGlySerSerValLeuSer 91
 Db 2772 -----TATTTGCTTCAATTCAAAATTAACGCCATATTGGAATTAACA 2813
 QY 92 GluAspLysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThr 111
 Db 92 ----- 111

Db 2814 GAAAACGACGTTAAAAATTAGAAAAG-----AATAATACAGTTCTTAATATCAATGCG-- 2867
 QY 112 GlyLysValCysAsnProAspAsnProGlnGluCysLeuLeuGluProGlyLeuAsn 131
 Db 2868 -----CCTGTTTCACT 2879
 QY 132 GluIleMetAlaAsnSerLeuAspTyrAsnGlnArgLeuTyrPalatPrgLysTyrParg 151
 Db 2880 GAAATA-----TATAGTGAGATATATTT----- 2903
 QY 152 SerGluValGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGlu 171
 Db 2904 -----GGAATTCATTAATAAGCT--TATGAGATTTTCCGCTGATTAAGTCCATA 2951
 QY 172 MetAlaArgAlaAsnHisTyrCysLysAspTyrGlyAspTyrPargGlyAspTyrGluVal 191
 Db 2952 TTAGGATCTTTTAATGCGCGTCCACACTTTTGGAAAGTTT--ACGGAAATTTCAATATA 3008
 QY 192 AsnGlyValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGlnHisThr 211
 Db 3009 AAG-----AAAAAATCATTAACAAAAGAAATAGTGCAT-- 3044
 QY 212 PheGluGluIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMet 231
 Db 3045 -----CATTAACAATTAATACATGAAATGAATATGCTTTA 3077
 QY 232 AsnAlaTyrProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuGluLysAsp 251
 Db 3078 GAAATTAAGCAATTAATGAAGAGTCTCTTAACCTCAAGAAATGTAATTTTGAATAT 3137
 QY 252 MetTyrGlyArgPheThrPheAsnLeuTyrSerLeuThrValProPheGlyGlnLysPro 271
 Db 3138 -----AATGAAATATATATACTTGTAAATTAATTAATTTACCT-----AAAAGT 3182
 QY 272 AsnIleAspValThrAspAlaMetValAspGlnAlaIleTyrPasnAlaGlnArgIlePheLys 291
 Db 3183 GATATAGATTAATTAATGACATATTATTGAGCTACATTAACAACTTATCATATATCT 3242
 QY 292 Glu-----AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnGly 308
 Db 3243 GAAAACATGATTAAGAAATGATTCCTGATCTATTCATGATTAATTAATGAGTTTGA 3302
 QY 309 -----PheTyrGluAsnSerMetLeuThrAspPro 318
 Db 3303 TCTGATTAATACAAATTTTAAAGAAATTCATTTGAAAACAAATTTATACAACT 3362
 QY 319 GlyAsn-----ValGlnLysAla 324
 Db 3363 ATAACTGAAGAGCAATTCATCTACCTTTTGTCCAGAAATTTATTAATAAAATAT 3422
 QY 325 ValCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIleLeuMetCysThr 344
 Db 3423 ATTTTAAACCAACGACACTTGGAAATTAATTCAGAAATAGTTTTCGA-----ACT 3473
 QY 345 LysValIleThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyHisIleGlnTyrAsp 364
 Db 3474 GAAATTTGGTTAAATGAGTTGCCAATATTT--AGAGAAATATGGCATATTTCCAAAAGAT 3530
 QY 365 MetAlaTyrAla----- 368
 Db 3531 GTAAATTAATGCTTTTGGAGATTAATGCTATTAATTTATTAATAATGACACAGCCTC 3590
 QY 369 -----AlaGlnProPheLeuLeuArgAsnGlyAla 378
 Db 3591 ATTATACATAAAAAAGACATAAAAAACATGCTAGATTCGAATTTTGAAGAGTTTATC 3650
 QY 379 AsnGluGlyPheHisGluAlaValGlyGluIleMetSer----- 391
 Db 3651 AATGATCTTAATAAAGAGAGATTTGAATAATTTAGCCATATATATAAAAAACTAGT 3710
 QY 392 -----LeuSerAlaAlaThrProLys-----HisLeuLys 401
 Db 3711 TTAAGAAGACATCTTTTCAATTATACCTAAAAATAGAAATTAAGACACTTCAATATATCTTAA 3770

CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Brookes, A. Anders
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB340P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 45:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 11384 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-527-45

Alignment Scores:
 Pred. No.: 0.00837 Length: 11384
 Score: 120.50 Matches: 146
 Percent Similarity: 31.28% Conservative: 108
 Best Local Similarity: 17.98% Mismatches: 238
 Query Match: 3.11% Indels: 321
 DB: 4 Gaps: 45

US-09-978-385-2_COPY_19_738 (1-720) x US-08-961-527-45 (1-11384)

QY 5 GUGUAGLALALSTHRPHELEUASRPLYSRPHASNHISGLUALAGLUSRLEUPHETRYGIN 24
 DB 11096 GAAGCATTAATAAACTTTATG-----ACGGGTGAAAATTTTATTCSCAA 11052
 QY 25 SerSerLeuAlaSer----- 29
 DB 11051 CATTACTAGACACATAGGAAGAATAATGAGAGCATGCTATACCTTCGCTTT 10992
 QY 30 TTPASNTYRASN-----ThrsnlethrGlu 39
 DB 10991 TGGGACACCTTAATCTCAGGCTGTTCACCTTGTTGGTGTGATTTTCCAACTGGATGGAAT 10932
 QY 40 AsnValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuYsgLugInser 59
 DB 10931 CAGATTCATGTAAGAAATGATTTTGGGCTCGGAGAGCTCTT-----ACC 10884
 QY 60 ThrLeuAlaGln-----MetLysProLeuGlnGluLeuGlnAsnLeuThrValYls 76
 DB 10883 AATATGGCTCAGAAAGGCGATATTAC----- 10857
 QY 77 LeuGlnLeuGlnAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 96
 DB 10856 ---AATATCATGCTCAGACCTCAAAATGGTCATCAATCATGATGAG----- 10815
 QY 97 ArgLeuAsnThrLeuAsnThrMetSerThrLeuLysSerThrGlnLysValCysAsn 116
 DB 10814 -----ATGACCTTTTTCGTCGATGATGAGCT----- 10785
 QY 117 ProAspAsnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 136
 DB 10784 -----CTGCCAGAAACAGGCGCAATCGTAACA----- 10758
 QY 137 SerLeuAspLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 153
 DB 10757 -----GAGCTTCCTGGAAGAAATGAGAGATGAGCTTTGGCTGGCAGAGAAAGAGCT 10704
 QY 154 ValGlyLysGlnLeuAspProLeuLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 173
 DB 10703 TGGGGCTTGAAGAGGCTCTGTC----- 10680
 QY 174 ArgAlaAsnHisLysLysAspLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 193
 DB 10679 -----AATATTTATGAAAGTTCAAGCTGATCATGGAAG-----AGAAAT 10641

QY 194 ValAspGlyTyrAspLysSerArgLysGlnLeuLeuGlnAspValGlnHisThrPheGlu 213
 DB 10640 TCTGATGGCAGTCCTTAATGCTTTGGCCAGAGATGAATGATCATCTTATCTGCTT 10581
 QY 214 GluIleLysProLeuThrGlnHisLeuHisAlaTyrValAlaGlnLysLeuMetAsnAla 233
 DB 10580 GAAATGAAAC-----TATTCCTCATATGAG----- 10557
 QY 234 TyrProSerLysLeuSerProLysLysLysLysLysLysLysLysLysLysLysLys 253
 DB 10556 -----TTATATGCC-----TGATGTCCTCATCTTTGGCTTATGATTTG 10518
 QY 254 GlyArgPheTrpThrAsnLeuLysSerLeuThrValProPheGlnLysPro----- 271
 DB 10517 GGGTATCAGCTTATGGGTACTTCGCTTATAGCATGCTTATGGCCAGACAGAGAGTTT 10458
 QY 272 -----AsnIleAspValThrAspAlaMetValAsp 281
 DB 10457 CAAGATTTTTCGAGAGAGTCATACCCATTAATATGGGCTT-----ATTGTGAC 10407
 QY 282 GlnAlaTrpAspAlaGlnArgLysPheLysGlnAlaGlnLysPhePheValSerValGly 301
 DB 10406 -----TGGTACACAGTCACT 10392
 QY 302 LeuProAsn-----MetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGly 319
 DB 10391 TTACCATCAACGATGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 10333
 QY 320 AsnValGlnLysAlaValCysHisProThrAlaTrp-----AspLeuGly 334
 DB 10332 GACCATTAATAGGCT-----CATACCAAGTGGTGGGTGCCCTTAATTTGACCTTGA 10279
 QY 335 LysGlyAspPheArg-----IleLeuMetCysThrLysValThrMetAspPhe----- 351
 DB 10278 AAAATGAAGATCAGATCCCTGTTATTTCTTCAATTAAGCATGATGATGATGATGATGAT 10219
 QY 352 -----LeuThrAlaHisHisGlnMetGlnHisGlnGlnGlnGlnGlnGlnGlnGln 366
 DB 10218 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10159
 QY 367 TyrAlaAlaGlnProPheLeu-----LeuArgAsnGlnAlaAsn-----GluGlyPhe 382
 DB 10158 -----CCATGACACCTTAATGAAGTGGCGGAATCTCACTAATGAAGTGAAT 10111
 QY 383 His-----GlnAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 10110 TATTTCCCTTACAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 10072
 QY 401 LysSerLysGlyLeuLeuSerProAspPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 420
 DB 10071 -----TATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10030
 QY 421 LeuLeuLysGlnAlaLeuThrLeuValGlyThrLeuProPheThrLysMetLeuGlnLys 440
 DB 10029 ATCAAGATTTACGGGATGAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 9979
 QY 441 TrpArg-----TrpMet-----ValPheLysGlyLysLeuPro----- 451
 DB 9978 TGAACATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9919
 QY 452 -----LysAspGlnTrpMet 456
 DB 9918 AATATGACTTTAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9859
 QY 457 LysLysTrpTrpGlnMetLysArgLysLysLysLysLysLysLysLysLysLysLys 476
 DB 9858 TTG-----CCATTCTCCGACGAT 9841
 QY 477 GluThrTyrCysAspProAlaSerLeuPheHisValSer-----AsnAsp 491
 DB 9840 GAAGTGGTTCATGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 9781
 QY 492 TyrSerPheIleArgTyrTyrThrArgThrLeuLysGlnGlnGlnGlnGlnGlnGlnGln 511

```

Db      9780 TTCGAGGCTT-----CGCAATCTTACCTACCA-----ATT 9745
QY      512 CysGlnAlaAlaLysHis----- 517
Db      9744 TGTACCCCTGTAAGAAATGTGCTTCATGGGTACGGAATACGGTCATTCCTAGAAATGG 9665
QY      518 -----GluGlyProLeu----- 521
Db      9684 AAATCTGAAGACAGTGGTAATGTCTTAACCTAGAAAGCCCAATGAATGTAAGATGAAG 9625
QY      522 -----HisLysCys----- 524
Db      9624 TATTTCGCTTCACCTAACCACTTTTACAAAGATCATCGCTGCTCTGTGGAAATTCAT 9565
QY      525 -----AspIleSerAsnSerThrGlnAlaGlnLysLeuPheAsn 538
Db      9564 ACCAGCTAATGATGATTAATGAATCATGTGAGGGGAATATGACAGACCAAGATGTTCTTCC 9505
QY      538 MetLeuArgLeuGlyLysSerLeuProThrPheLeuAlaLeuGlnAsnValValGlyAla 558
Db      9504 TTTATTCTGTAAGGTAAGAAAGGAGAAATGTTAGTCTGTATCTTAATATGTAACCTGTT 9445
QY      559 LysAsnMetAsnValArg-----ProLeuLeuAsnTyrPheGluProLeuPhe--- 574
Db      9444 GAGCGGAAAGATTTTACATCGGACTACCCGTTGACAGAAATTTACGAAAGATATGAAAT 9385
QY      575 ThrTrpLeuLysAspAsnLysAsnSerPheValGlyTrpSerThrAspTrpSerPro 594
Db      9384 ACTGATGTGGAAGAG-----TGGGAGGCGCTTGGAAAGAA 9349
QY      595 TyrAlaAspGlnSerIleLysValArgIleSerLeuLysSerAlaLeuGlyAspLysAla 614
Db      9348 CAT---AAATCAACGCTTCAACGCAAGAGACTA----- 9316
QY      615 TyrGluTrpAsnAspAsnGlnMetTyrLeuPheArgSerSerValAlaTyrAlaMetArg 634
Db      9315 -----TGGAAAGATTATGACGACACTTAACCTTACCCTACCGCTATGGGACCAAGT 9262
QY      635 GlnTyrPheLeuLys-----ValLysAsnGlnMetIleLeuPheGlyGlnGluAspVal 652
Db      9261 GATATGAAATCAACGCTGCTTGAATCTACTAATAACCGTCACAAATAAACAACAAAA 9202
QY      653 ArgValAlaAsnLeuLysProArgIleSerPheAsn 664
Db      9201 GGAGTAGAAATGAATAATGTAATGTAGCTTGTAT 9166

RESULT 17
US-08-157-171-4
; Sequence 4, Application US/08157171
; Patent No. 5736323
; GENERAL INFORMATION:
; APPLICANT: Soubrlier, Florent
; APPLICANT: Hubert, Christine
; APPLICANT: Corvol, Pierre
; TITLE OF INVENTION: Agents and Procedures for the Study of
; TITLE OF INVENTION: The Genetic Polymorphism of the Angiotensin I Converting
; TITLE OF INVENTION: Enzyme
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 90 South 7th Street, 3100 No. 5736323west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/157,171

```

```

; FILING DATE: 24-JAN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hillson, Randall A.
; REGISTRATION NUMBER: 31,838
; REFERENCE/DOCKET NUMBER: 8076.103USMO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Exon 17
; US-08-157-171-4

Alignment Scores:
Pred. No.: 5.87e-06 Length: 144
Score: 120.00 Matches: 21
Percent Similarity: 59.57% Conservative: 7
Best Local Similarity: 44.68% Mismatches: 19
Query Match: 3.10% Indels: 0
Gaps: 0

US-09-978-385-2_copy_19_738 (1-720) x US-08-157-171-4 (1-144)

QY      130 LeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGluArgLeuTrpAlaTrpGlnSer 149
Db      3 CTGACGAATGTGATGGCCACATCCCGAAATTAAGAACTGTATAGGCGATGGAGGCG 62
QY      150 TrpArgSerGluValGlyLysGlnLeuArgProLeuTyrGlnGluTyrValLeuLys 169
Db      63 TGGGACACAAAGGGGGGAGAGAGCATCCCTCACTTTTACCCGAAATACGTGGAACTCATC 122
QY      170 AsnGluMetAlaArgAlaAsn 176
Db      123 AACGAGCTGCCGCGCTCAAT 143

RESULT 18
US-08-961-527-115/C
; Sequence 115, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1

```


Db 1534 -----TAAAGTTGG----- 1526
 QY 579 AspGlnAsnLysAsnSerPheValGlyTyrSerThrAspTyrSerProTyrAla----- 596
 Db 1525 -----TGGAGTCATGATATGCGTGGACATATGTTGGGG 1493
 QY 597 -----AspGlnSerIleLysValArg-----IleSerLeuLysSer 608
 Db 1492 GATTATGACGAGAGCTACAACTGATGCGTACAGTTTGCTTTTATAGTCCGATTTACT 1433
 QY 609 AlaLeuGlyAspLysAlaTyrGluTyrPheAsnAspAsnGluMetTyrLeuPheArgSerSer 628
 Db 1432 CGATTACATGATGTTAGAGTCTTTTATAGTAAAGAACCTTGTTTTCACAGAACCA 1373
 QY 629 ValAlaTyrAlaMetArgIleTyrPheLeuLysValLysAsnGluMetIle 645
 Db 1372 ACATCTAGATTTATGAGAAATATC---CTTCGTTTGAGACATCAATGATTT 1325

RESULT 19

US-08-645-193B-18
 ; Sequence 18, Application US/08645193B
 ; Patent No. 5962253

GENERAL INFORMATION:

; APPLICANT: Kupke, Thomas
 ; APPLICANT: Goetz, Friedrich
 ; APPLICANT: Kemper, Christoph
 ; APPLICANT: Jung, Gunther
 ; TITLE OF INVENTION: Oxidative Decarboxylation of Peptides
 ; TITLE OF INVENTION: Catalyzed by Flavoprotein Epld
 ; NUMBER OF SEQUENCES: 70
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 ; STREET: 1100 New York Avenue, Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/645,193B
 ; FILING DATE: 13-MAY-1996

CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
 ; NAME: Esmond, Robert W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 0652,1540000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600
 ; TELEFAX: (202) 371-2540
 ; INFORMATION FOR SEQ. ID NO: 18:

SEQUENCE CHARACTERISTICS:

; LENGTH: 8700 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: both
 ; MOLECULE TYPE: cDNA
 ; US-08-645-193B-18

Alignment Scores:

Pred. No.: 0.00804 Length: 8700
 Score: 119.00 Matches: 150
 Percent Similarity: 32.248 Conservative: 115
 Best Local Similarity: 18.25% Mismatches: 247
 Query Match: 3.08% Indels: 310
 DB: 2 Gaps: 40

US-09-978-385-2_COPY_19_738 (1-720) x US-08-645-193B-18 (1-8700)

QY 2 ThrIleGluGluAlaLysTyrPheLeuAspLysPheAsnIleGlu----- 17
 Db 2619 ACTATATGGAATTTATCAGCAATTTTATGCAATAATATGATTTGAACAACATGATAAT 2678
 QY 18 -----AlaGluAspLeu-----PheGlyClnSerSerLeuAlaSerPheAsn 31
 Db 2679 TTAAGCAATTCCTCTGACATATATAATGATTTGGCTATCCAAAAAGACACTTATAGT 2738
 QY 32 TyrAsnThrAsnIleThrGluLysValGlnAsnMetAsnAlaGlyAspLysTyr 51
 Db 2739 TTTTCTAATATCAAT----- 2753
 QY 52 SerAlaPheLeuLysGluGlnSerThrLeuAlaGlnMetLysProLeuGlnGluIleGln 71
 Db 2754 ---CCATTTTAAAGAAAG----- 2771
 QY 72 AsnLeuThrValLysLeuGlnIleGlnAlaLeuGlnIleAsnGlySerSerValLeuSer 91
 Db 2772 -----TATTTGCTTGCATTAATCAAAATATACAGCCATATTTGAATTAACA 2813
 QY 92 GluAspLysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThr 111
 Db 2814 GAAACGACGCTTAAATTTTGAAGAG---AATATATCAGTTCTTAAATCAATGCG--- 2867
 QY 112 GlyLysValLysAsnProAspAsnProGlnGluCysLeuLeuGluProGlyLeuAsn 131
 Db 2868 -----CCTGTTTCACAT 2879
 QY 132 GluIleMetAlaAsnSerLeuAspTyrAsnGluArgLeuThrAlaTyrGluSerTyrPhe 151
 Db 2880 GAAATTA-----TATAGGAGATATATATTT----- 2903
 QY 152 SerGluValGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGlu 171
 Db 2904 -----GGAATTCATATAAAGG---TATGAGATTTTCCCGGATTAATCCAAAT 2951
 QY 172 MetAlaArgAlaAsnIleTyrGluAspTyrGlyAspTyrTrpArgLysPyrGluVal 191
 Db 2952 TTAGGACCTTTTATGCGCGGTGCAACTTTTGAAGGTTT---ACGGAAATTTCAATATA 3008
 QY 192 AsnGlyValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGluIleThr 211
 Db 3009 AAG-----AAAAAATCAATTTACAAAAGCAATATGCGCAT--- 3044
 QY 212 PheGluGluLysProLeuTyrGluIleAsnIleAlaTyrValArgAlaLysLeuMet 231
 Db 3045 -----CATTAACATATATATCAATGAATGAATATGCTTA 3077
 QY 232 AsnAlaTyrProSerTyrIleSerProIleGlyCysLeuProAlaIleLeuGlyAsp 251
 Db 3078 GAAATTAAGCCAAATTAATGAAGGCTCTTAACCTCAAGAAATGTAATATTTGAATTAAT 3137
 QY 252 MetTrpGlyArgPheTrpThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysPro 271
 Db 3138 -----AATAGCAATATATATCTGTTTAAATTTAAATTTACT-----AAAAGT 3182
 QY 272 AsnIleAspValThrAspAlaMetValAspGlnAlaThrAspAlaGlnArgIlePheLys 291
 Db 3183 GATATAGATATAAATGACATATTTATGAGAGCTACATTTAAACAACCTTATCATATATCT 3242
 QY 292 Glu-----AlaGluLysPhePheValSerValGlyLeuProAsnMetIleGlnGly 308
 Db 3243 GAAAAACATGATTCAGAAATTTGATGCTATCTAATTCATATGTTAATTTGAGATTGGA 3302
 QY 309 -----PheTrpGluAsnSerMetLeuThrAspPro 318
 Db 3303 TCTGAATTTATCAATATTTTAAAGAAATTTTCATTTGAAAAACAAATTTTATACACCT 3362
 QY 319 GlyAsn-----ValGlnLysAla 324
 Db 3363 ATTAAGTGAAGAGCATTCATTCACCTTTTGTCCAGAAATTAATTAATAAATATAT 3422
 QY 325 ValCysHisProThrAlaThrPheAspLeuGlyLysGlyAspPheArgIleLeuMetCysThr 344


```

Db      3423 ATTTTAAACGACGCTACTGGAAAAATTAATTCAGAAATGTTTCTGAA-----ACT 3473
QY      345 LysValThrMetAspSerPheLeuThrAlaHisHisGluMetCysHisIleGlnTyrAsp 364
Db      3474 GAAAATGGTTAAATGTTGCGAACTATT---AGAGAAAAATGGCAATATTCGAAAAGAT 3530
QY      365 MetAlaTyrAla-----
Db      3531 GATATTAATGCTTTTGGAGATATGATGCTATTAAATTTAATTAATGACAAACATCTTC 3590
QY      369 -----AlaGlnProPheLeuLeuArgAsnGlyAla 378
Db      3591 ATTATACATAAAAAAGACTAAAAAACAATGGTAGATTCGAATATTAGAAAGCTTTATC 3650
QY      379 AsnGlnGlyPheHisGlnAlaValGlyGluIleMetSer----- 391
Db      3651 AATGAATCTTAATTAATAGAGAAATGTAGAAATTTGTACCCATTAATTAATAAACTACT 3710
QY      392 -----LeuSerAlaAlaThrProLys-----HisLeuLys 401
Db      3711 TTTAAAGAACAACTTTCATTAATACCTAAATAATAGAACATTCATTAATCTTAA 3770
QY      402 -----SerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIle 418
Db      3771 GATTGGTTTCAATTCATTAATTAATGATCTTAAACATACCAAGAT----- 3815
QY      419 AsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeu 438
Db      3816 AATTTTATTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3875
QY      439 GlnLysThrArgTyrMetValPheLysGlyGluIleProLysAspGlnTyrMetLys 458
Db      3876 AATAAATTTTAAACATAAATTTTAAAGAGAT-----GAGATTTTAAATAAA 3923
QY      459 ThrTrpGlnMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThr 478
Db      3924 ---TTAGATTATTAGAGAA-----GATGAGAGAT 3950
QY      479 TyrCysAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyr 498
Db      3951 TAT-----TTCCAATTAATTAATTCATTAATAAAATTTGG 3983
QY      499 -----ThrArgThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCys 512
Db      3984 AAGATTATTCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4043
QY      513 GlnAlaAlaLysHisGluGlyProLeuHisLysCysAspIleSerAsnSerThrGluAla 532
Db      4044 GAAATATATGATATGTTGGTGGCCACACGTAATGAGATATGAGATTTTATTAATGAT 4103
QY      533 GlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGlnProThrThrLeuAlaLeu 552
Db      4104 GATGCTGCTATTA-----TCAATTAATTAATTAATTAATTAATTAATTAATTAAT 4154
QY      553 GluAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGluPro 572
Db      4155 GAATTTATGCTTGTCT-----ATATCAATTAATTAATTAATTAATTAATTAATTA 4202
QY      573 LeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGlyTyrPserThrAspTyr 592
Db      4202 ----- 4202
QY      593 SerProTyrAlaAspGlnSerIleLysValArgIleSerLeuLysSerAlaLeuGlyAsp 612
Db      4203 -----ATTATTAATAAGTAGAGAAAGAAAGAAATTTAATTAATTAATGCG-----GAA 4247
QY      613 LysAlaTyrGlnTyrPheAsnAspAsnGluMetLysLeuPheArgSerSerValAlaTyrAla 632
Db      4248 GATTTAATATCGTACTAATGAC----- 4268
QY      633 MetArgGlnTyrPheLeuLysValLysAsnGlnMetIleLeuPheGlyGluGluAspVal 652

```

```

Db      4269 ATTAAGACATAT----- 4280
QY      653 ArgValAlaAsnLeuLysProArgIleSerPheAsnPhePheValThrAlaProLysAsn 672
Db      4281 -----AAAAATTTATGCTAAACT-----ACCAATCTCAAAAT 4316
QY      673 ValSerAspIleIlePro----- 678
Db      4317 GACTATGAATTTTAAAAAAGAAATTCGCAATCTCATGAAATTTAATTAATTAAT 4376
QY      679 -----ArgThrGluValGlnLysAlaIleArgMetSerArgSerArg 692
Db      4377 AGTATTTAGAAATCTTAAAAAGACACTACAAAAAGCTTAATTAATTAATTAATTA 4436
QY      693 IleAsnAspAlaPhe-----ArgLeuAsnAspAsnSerLeuGluPheLeuGlyIle 709
Db      4437 ATAATTTGCAGCTTTTATACACATGCGTTGTAT-----AGAAATTTGCGTAT 4484
QY      710 GlnPro 711
Db      4485 AATTCCT 4490

RESULT 20
US-08-630-916A-45
Sequence 45, Application US/08630916A
Patent No. 6011137
GENERAL INFORMATION:
APPLICANT: Pirozzi, Gregorio
APPLICANT: Kay, Brian K.
APPLICANT: Fowlkes, Dana M.
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,916A
FILING DATE: 03-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 896-8864/9741
INFORMATION FOR SEQ. ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 2052 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-630-916A-45

Alignment Scores:
Pred. No.: 0.0016 Length: 2052
Score: 116.00 Matches: 141
Percent Similarity: 32.17% Conservative: 107
Best Local Similarity: 18.29% Mismatches: 265
Query Match: 3.00% Gaps: 258
DB: 36

```

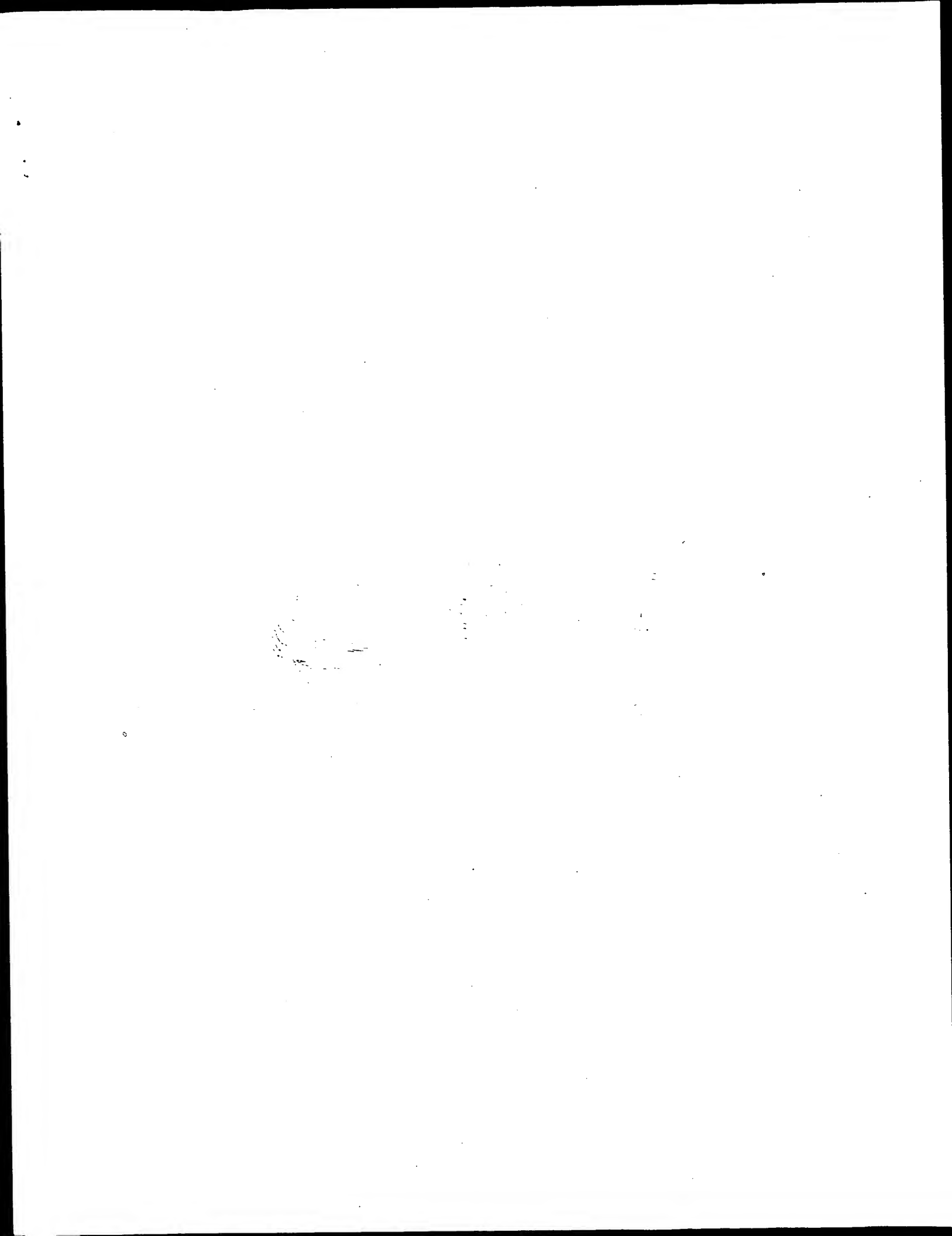
US-09-978-385-2_COPY_19_738 (1-720) x US-08-630-916A-45 (1-2052)

OY		12	AspLysrheksnhiisgiValaGluAspLeuPheTyrGlnSerSerLeuAlaSerTrpAsn	31
Db		152	GAACTGTATTGGAGATATCATCTCTCA---TTTGCCACCACTGAATAAAGGGTCGTGCACG	208
OY		32	TyrAsnThrIleThrglnGluAspValaGlnAsnMetAsnAlaGlyAspLysTrp	51
Db		209	GGTAGTCACACTAGTGTGTGAAGAANAATGCCCTGTCTCCAAATTGGCACT-----	256
OY		52	SerAlaPheLeuLysGlnGlnSerThrLeuAlaGlnMetCysProLeuGlnGluIle--	70
Db		257	-----AGTACTACHTGGAGATGATCCAGTCCAGTTCACAAGAAATACGTG	295
OY		71	-----GlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeuGlnGlnGlnAsnGly	86
Db		296	ACTTCCTCAGAAAACATGATATATATTCCTTACCACAGTCCACAAATTTGGATGTGAAGCT	355
OY		87	SerSerValLeuSerLysLysSerLysArgLysAsnThrIleLeuAsnThrMetSer	106
Db		356	AGAAATATTTATTAAGCTCGACCTCACCTCTMAAT-----TCTAAGAGNAGT	397
OY		107	ThrIleTyrSerThrGlyLysValLysAsnProAspAsnProGlnIucylSLeuLeu	126
Db		398	TCGTGCTTTTGAGCACCACCAAATTAACAACGCCAGATGGGTGTATGATCTGTACGGCAG	457
OY		127	GluProGlyLeuAsnGlnIleMetAlaAsnSer-----	137
Db		458	CAGTGTGGG-----AATGCCAACACAGAAACCTTGGCATCAGGGTGGGAACAA	505
OY		138	-----LeuAspTyrAsnGlnIucylArgLeuThrAlaTrp	147
Db		506	AGAAAAAGATCCCTCAGTGTGAACCTTATGTGGGACAAATAACCTGCAACTCCACATGG	565
OY		148	GlnSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyrGlnGluIucylVal	167
Db		566	GAGAGA-----CCACACCTTTACCTCCAGGTGGGAAAGAGAGTTGAT	610
OY		168	LeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrpArgLy	187
Db		611	GATCGTAGAAGATTATTTATGTGATCTCAATACCCAGAACAAACCTGGCAGCGGCT	670
OY		188	AspTyrGlu---ValAsnGlyValAspGlyTyrAspTyrSerArgLysGlnLeuLeu	206
Db		671	ACCAGGAATCTGTCCGAATTTTGAACAGTGGCAATCTCAGCGGACCAATGG-----	724
OY		207	AspValaGlnHisThrPheGlnGluIleLysProLeuTyrGlnHisLeuHisAlaTrpVal	226
Db		725	-----CAGGGAGCTATGACMACAGTTTAACCAACAGATCCTGTATGGCTTCAMTGTTA	778
OY		227	ArgAlaLysLeuMetAlaLayTrpProserTyrLieserProLigLysLeuProAla	246
Db		779	GCTGCAGAA-----AATGACCTTATGAGACCTTGTGCACCA	814
OY		247	HisLeuLeuGlyAspMetTrpGlyArgPheTrpHisLeuTyrSerLeuThrValPro	266
Db		814	-----	814
OY		267	PheGlyGlnLysProAsnIleAspValThrAspAlaMet-----	279
Db		815	-----GGCTGGGAAAAAAGATGATTCACACAGCAGGGTTTACTTTGTGATCATATACACA	871
OY		280	-----Valasp	281
Db		872	AAACAAACCCAGTGGAGATCCAAGACTCAAGCTTACAGAAATGAAGAAACCCCTGGCA	931
OY		282	GlnAlaTrpAspAlaGlnArgIlePheLysGlnAlaGluLysPhePheValSerValGly	301
Db		932	GAAAGCTGTGGGAAAT---AGATATATACCTGTAAGATGTGAAGTACTTTGTTGATCATATAC	988
OY		302	LeuProAsnMetThrGlnGlyPheTrpLysnSerMetLeuThrAspProGlyLysAla	321
Db		989	ACAAGAACACAAACA-----TTCAAGATCTCTGGCAATGGG	1024

OY	322	GLNLYALALALYSHS-----	ProthAlatrAspleuclYlysglyAspRne	338
		::: :::: :: ::	:: :: ::	
Db	1025	AAATCATCTGTAACTAAAGGTGGTCCACAAATTCCTTAAGAACGGGC-----	TTT	107
OY	339	ArgIleLeuMetCysThrIysValThrMetAspPheLeuThrAlaHisHsIglMet	358	
Db	1076	AGS-----	-TGAAGCTT	108
OY	359	GLYHisIleGlnItyrAspMetAlaItyrAlaAlaGlnProPheLeuLeuArgAsnIyala	378	
		:::: ::		
Db	1088	GCTCACTCCGTTAT-----		110
OY	379	AsnGlnIyPheHisGlnAlaValIgluIleMetSerLeuSerAlaIathrProlys	398	
Db	1103	-----	-TTGGCAGCTTAAATCAGCTCACT	112
OY	399	HisLeuLYSerIleGlyLeuLeuSerProAspPheGlnIAspAsnGlnItyrGluIle	418	
		:: :: :: ::	:: :: :: ::	::: ::
Db	1130	CATGTAAG-----ATCAATGTGCCCCGCGACAGACATTGTGGAAGATTCCTTCACAGATT	118	
OY	419	AsnPheLeuLeuLYsGlnAlaLeuThrIleValIglYthLeuProPheThrItyrMetLeu	438	
Db	1187	-----ATGGCATTTAAA-----	-CCCATGACTGG	121
OY	439	GluLYSerItyrArgTyrMetValPheLYsgIgluIleProLYsAspGln-----	-Trmet	456
		::: ::: :: :: :: :: :: ::		
Db	1211	ACGAGGCGCTTAATATGATATATTTAGAGAGAAAGACCTGATATAGTGCGCTACGC	127	
OY	457	LYsLYsItyrTrpIleMetLYsArgGluIleValIglValIgluProVal-----		473
		:: :: :: ::	:: :: ::	
Db	1271	ACGAAATGGATTTCTTGCTTCCACAT-----	GAAGTTTTCACCAATGATTTGCTTA	132
OY	474	-----ProHisAspGluItyrThrCys-----	AspProAlaSerLeuHe-----	486
Db	1325	TTTGTAGTATCGCGGGAAGAACACTATGTCTGCGAGATTAATCCAGCATCAACCTTAAT	138	
OY	487	-----HisValSerAsnAspTyrSerPheIle-----	ArgTyrTyrThrArgThrLeuTyr	503
		:: :: :: ::	:: :: ::	
Db	1385	CCAGACCAATCTTCATAC-----TTCTGTTTCATTTGGCGTTTATATGTCATGCACCTAATT	144	
OY	504	GlnPheGlnPheGlnIgluAlaLeuLYsGlnAlaAlaLYsHisGluIgluProLeuHisLYs	523	
		:: :: :: ::	:: :: ::	
Db	1442	CATGAAAGTTTATGATACACGCTGTTCTTTCATTCACCTTACACACGCTATGTTAAGTAA	150	
OY	524	-----CysAspIle-----		526
Db	1502	AAACTTACTATTAAAGATTTTGAATCTATGATACTGAATTTTAACTCCCTTATCTCG	156	
OY	527	-----SerAsnSerThrGlnAlaIgluIlyLeuPhe-----	AsnMetLeu	540
		:: :: :: ::	:: :: ::	
Db	1562	ATTAAGACATTAACACATTTGAAGAAATGTCGTTAGAAATGATCTTCTGTTGTCACATGAG	162	
OY	541	ArgLeuGluIlyLYsSerGlnProTyrThrIleuAlaLeuGlnAsnValIAlaIlyAsn	560	
		:: :: :: ::	:: :: ::	
Db	1622	ATTTTGGGAAAGTACTTCACTACATGACCTGAAGTTG-----	CGAGGTTCAT	166
OY	561	MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrItyrLeuLYsAspGln	580	
Db	1670	ATTCTGGTG-----	-ACTGAGGAG	168
OY	581	AsnLYsAsnSerPheValIglYTrpSerThrAspTyrSerProLYsAlaAspGlnSerIle	600	
		:: :: :: ::	:: :: ::	
Db	1688	AACAAGATGATATATTTGTTTATATGACAGAAATGG-----		1723
OY	601	LYsValArgIleSerLeuLYsSerAlaLeuGlyAspLYsAlaItyrGluItyrPAsnAspAsn	620	
Db	1724	-----CGTTTCTTCAGAGAGTCAAGAACAGACCAAGAGCTTCCCTGATGCTTTAAT	177	
OY	621	GluMetItyrLeuPheArgSerSerValAlaItyrAlaMetArgGlnItyrPheLeuLYsVal	640	
		:: :: :: ::	:: :: ::	:::
Db	1778	GAAGTT-----	-GTTCCTTCACAGTGGCTACAGTCTTCGATGAGAAAA	1819

QY 641 LysAsnGlnMetIleLeuPheGlyGluAspValArgValAlaAsnLeuLysProArg 660
Db 1820 GAATTAGAGGTTATGTTGTGCGATGCAGAGTTGACTTGCAGATTGGCAGAGAAAT 1879
QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSer-----AspIleIle 677
Db 1880 ACTGTTTATCGACATTATACAAAGAAACAGCAATCATTTGTTTGGCAGTTTGTG 1939
QY 678 ProArgThrGluValGluLysAlaIleArgMet 688
Db 1940 AAAGAGCAGACAAATGAGTAAGAAATGCGACTA 1972

Search completed: May 26, 2003, 20:22:58
Job time : 150 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 18:06:24 ; Search time 268 Seconds
(without alignments)

3547.511 Million cell updates/sec

Title: US-09-978-385-2_COPY_19_738

Perfect score: 3869
Sequence: 1 STIEBQAKFLDKFHEAD.....DNSLFLGIQPTLGPNNPP 720

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frimex_p2n model -DEV=11h
-O=/cg2_1/USPTO.spool/US09978385/runat_21052003_154420_19915/app.query.fasta_1.903
-DB=Published.Applications_NA -OFMT=fastap -SUFFIX=trpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pcr -THR_MAX=100
-THR_MIN=0 -ALIGN=20 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09978385 -ECGN_1_1_57=runat_21052003_154420_19915
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEOUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published.Applications_NA:*

1: /cg2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/ptodata/1/pubpna/US06_PUB.seq:*
3: /cg2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cg2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cg2_6/ptodata/1/pubpna/PCUS_PUBCOMB.seq:*
7: /cg2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cg2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cg2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cg2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cg2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cg2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	3334	9	US-09-978-385-1
2	3869	100.0	3369	9	US-10-158-847-141
3	3865	99.9	3325	9	US-10-114-893-85
4	3833	99.1	2920	9	US-09-969-384-2

5	3833	99.1	2920	9	US-10-158-847-137	Sequence 137, App
6	3785	97.8	2911	9	US-09-969-384-12	Sequence 12, App1
7	3720	96.1	3732	9	US-10-028-072-71	Sequence 71, App1
8	3720	96.1	3732	9	US-10-121-049-71	Sequence 71, App1
9	3720	96.1	3732	9	US-10-123-904-71	Sequence 71, App1
10	3720	96.1	3732	9	US-10-140-470-71	Sequence 71, App1
11	3720	96.1	3732	9	US-10-175-746-71	Sequence 71, App1
12	3720	96.1	3732	9	US-10-176-918-71	Sequence 71, App1
13	3720	96.1	3732	9	US-10-176-921-71	Sequence 71, App1
14	3720	96.1	3732	9	US-10-137-865-71	Sequence 71, App1
15	3720	96.1	3732	9	US-10-140-474-71	Sequence 71, App1
16	3720	96.1	3732	9	US-10-142-431-71	Sequence 71, App1
17	3720	96.1	3732	9	US-10-143-114-71	Sequence 71, App1
18	3720	96.1	3732	9	US-10-140-002-71	Sequence 71, App1
19	3720	96.1	3732	9	US-10-142-419-71	Sequence 71, App1
20	3720	96.1	3732	9	US-10-123-262-71	Sequence 71, App1
21	3720	96.1	3732	9	US-10-142-423-71	Sequence 71, App1
22	3720	96.1	3732	9	US-10-121-050-71	Sequence 71, App1
23	3720	96.1	3732	9	US-10-141-755-71	Sequence 71, App1
24	3720	96.1	3732	9	US-10-143-032-71	Sequence 71, App1
25	3720	96.1	3732	9	US-10-123-108-71	Sequence 71, App1
26	3720	96.1	3732	9	US-10-123-236-71	Sequence 71, App1
27	3720	96.1	3732	9	US-10-123-261-71	Sequence 71, App1
28	3720	96.1	3732	9	US-10-140-921-71	Sequence 71, App1
29	3720	96.1	3732	9	US-10-140-928-71	Sequence 71, App1
30	3720	96.1	3732	9	US-10-121-045-71	Sequence 71, App1
31	3720	96.1	3732	9	US-10-123-292-71	Sequence 71, App1
32	3720	96.1	3732	9	US-10-123-903-71	Sequence 71, App1
33	3720	96.1	3732	9	US-10-124-819-71	Sequence 71, App1
34	3720	96.1	3732	9	US-10-124-822-71	Sequence 71, App1
35	3720	96.1	3732	9	US-10-140-925-71	Sequence 71, App1
36	3720	96.1	3732	9	US-10-160-498-71	Sequence 71, App1
37	3720	96.1	3732	9	US-10-121-041-71	Sequence 71, App1
38	3720	96.1	3732	9	US-10-121-043-71	Sequence 71, App1
39	3720	96.1	3732	9	US-10-121-047-71	Sequence 71, App1
40	3720	96.1	3732	9	US-10-123-215-71	Sequence 71, App1
41	3720	96.1	3732	9	US-10-123-902-71	Sequence 71, App1
42	3720	96.1	3732	9	US-10-123-908-71	Sequence 71, App1
43	3720	96.1	3732	9	US-10-123-909-71	Sequence 71, App1
44	3720	96.1	3732	9	US-10-123-910-71	Sequence 71, App1
45	3720	96.1	3732	9	US-10-124-813-71	Sequence 71, App1

ALIGNMENTS

RESULT 1
US-09-978-385-1
Sequence 1, Application US/09978385
Patent No. US20020177211A1
GENERAL INFORMATION:
APPLICANT: Piddington, Christopher S.
APPLICANT: Petrie, Charles
APPLICANT: Shoemaker, Kimberly E.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME
FILE REFERENCE: 99-24C1
CURRENT FILING DATE: 2001-10-16
PRIORITY FILING DATE: 1999-05-13
PRIORITY APPLICATION NUMBER: 60/133,952
PRIORITY FILING DATE: 1999-08-27
PRIORITY APPLICATION NUMBER: 09/563,516
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 3334
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (35)...(2449)

US-09-978-385-1

Alignment Scores:

Pred. No.:	0	Length:	3334
Score:	3869.00	Matches:	720
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
		Gaps:	0

US-09-978-385-2_COPY_19_738 (1-720) x US-09-978-385-1 (1-3334)

QY 1 SerThrIleGluGluGlnIleAlaIleYsthrPheLeuaspPysPheAsnHisGluAlaGluasp 20
 Db 89 TCACCATTTGAGGAGACAGGAGCCAGATTTTGGACAGATTAAACCAAGCCGAGGAGAC 148
 QY 21 LeuPheTyrglnSerSerIleuAlaSerTrpAsnThrAsnIleThrGluGluAsn 40
 Db 149 CTGTCTCATCAAGTTCACCTGCTCTTGGAAATTATAACACCATATTACTGAGAGAAAT 208
 QY 41 ValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuYsgluGlnSerThr 60
 Db 209 GTCCAAACATGAATTAATGCTGGGAGCAAAATGCTCTCTTTTAAAGAAACAGTCCACA 268
 QY 61 LeuAlaGlnMetTyrrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 Db 269 CTGCCCAAAATGATCCACATCAAGAAATTCAGAAATCTCACAGTCAAGCTTCAGCTGACG 328
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 Db 329 GCTCTTCAGCAAAATGGGCTTCAGTCTCTCAGAAACAGCAAGCAAAAGCTTGAACACA 388
 QY 101 IleLeuAsnThrMetSerThrIleIleTyrrSerThrGlyLysValLysAsnProAspAsnPro 120
 Db 389 ATTTTAATATCAATGACACCATCTACAGTACAGTGAAGATTGTTAAACCAAGATATCCCA 448
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrr 140
 Db 449 CAGAAATGCTTATTACTTGAACACAGCTTGAATGAATATATGGCAAAAGCTTATGACTAC 508
 QY 141 AsnGluArgLeuTrpAlaIleTrpGluSerTrpArgSerGluValLysGlnLeuArgPro 160
 Db 509 AATGAGAGGCTCGGGGCTGGGAAGCTGAGATCTGAGCTGGCAGCAAGCTGAGGCCA 568
 QY 161 LeuTyrglnGluIleTyrrValIleLeuLysAsnGluMetAlaArgAlaAsnHisTyrrGluAsp 180
 Db 569 TTATATGAAAGATATGGGTCTTGAATAAATGAGATGGCAGACCAATCATTTATGAGGAC 628
 QY 181 TyrrGlyAspTyrrTrpArgGlyAspTyrrGluValAsnGlyValAspGlyTyrrAspTyrrSer 200
 Db 629 TATGGGATTAATGGAGAGAGACTATGAAGTAAATGGGGTAAATGCTATGACTACACG 688
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrrGlu 220
 Db 689 CGGGGCCAGTTGATTAAGATGGGAAACATACCTTTGAAGAGATTTAAACCATTAATATGAA 748
 QY 221 HisLeuHisAlaTyrrValArgAlaLysLeuMetAsnAlaTyrrProSerTyrrIleSerPro 240
 Db 749 CACTTCATGAGCCATATGAGGGGCAAAAGTTGATGATGCTTCTTCAATATCCAGTCCA 808
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeu 260
 Db 809 ATTGAGATGGCTCTCTCTCATTTGCTTGGTATATGTTGGGTAGATTTGGCAAAATGCG 868
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 Db 869 TACTCTTTGACAGTTCCTTTTGGACAGAAACCAACATAGATGTACTGATCCAAATGGTG 928
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 Db 929 GACCAAGGCTGGGATTCACAGAGAAATATTCAGAGAGGCCGAGAAAGTTCTTTATCTCTG 988
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyLysn 320

Db 989 GGCTTCCTATATGACTCAAGCATTCGGGAAATTTCCATGCTTAACGAGCCAGGAAAT 1048
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 Db 1049 GTTCAGAAACACAGTCCCATCCACAGCTTGGGAGCTGGGAGAGGCCAGCTTCAGAGATC 1108
 QY 341 LeuMetCysThrLysValThrMetAspPhePheLeuThrAlaHisIleGluMetGlyHis 360
 Db 1109 CTTATGTGCAAAAGGTGACATGAGAGACTTCTGACAGCTCATCATATGATGGGGCAT 1168
 QY 361 IleGlnTyrrAspMetAlaTyrrAlaAlaGlnProPheLeuLeuArgAsnGlyLysAsnGlu 380
 Db 1169 ATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1228
 QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 Db 1229 GGATTCACAGAACTTGGGGAATCAATGTCATCTTTCGACGACACACCTTAAGCATTTTA 1288
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluLeuAsnPhe 420
 Db 1289 AATTCATTTGCTTCTGTCACCCGATTTTCAAGAAAGCAATGAAACGAAATTAACATTC 1348
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrrMetLeuGluLys 440
 Db 1349 CTGCTCAAAACAGCACTCAGATGTTGGGACTGCGCATTTACTTACATGTTACAGAAAG 1408
 QY 441 TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpArg 460
 Db 1409 TGGAGGTGATGTGCTTTTAAAGGGGAAATTTCCAAAGCAACATGATGATGATGATGAT 1468
 QY 461 GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrrCys 480
 Db 1469 GAGATGAACGAGAGATAGTGGGGTGGGAACTGGGCCCATGATGAAATATACGTGT 1528
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrrSerPheIleArgTyrrThrArg 500
 Db 1529 GACCCCGCATCTCTCTTCATGTTCTTAATGATTAATCATTCATTCATTCATTCATTCAT 1588
 QY 501 ThrLeuTyrrGlnPheGlnIlePheGlnIleAlaLeuCysGlnIleAlaLysGluGlyPro 520
 Db 1589 ACCCTTTCACATTCACATTCACATTCACATTCACATTCACATTCACATTCACATTCACAT 1648
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
 Db 1649 CTGCAAAATGTGACATCTCAAACTCTACAGAACTGAGAGAAACGTTCAATATGCTG 1708
 QY 541 ArgLeuGlyLysSerGluProTrpPheLeuAlaLeuGluAsnValValGlyAlaLysAsn 560
 Db 1709 AGCTTGGAAATTCAGAAACCTGGACCTTACCATTTGAAATGTTGTAAGGAGCAAAAGAAC 1768
 QY 561 MetAsnValArgProLeuLeuAsnTyrrPheGluProLeuPheThrTrpPheLysAspGln 580
 Db 1769 ATGATGTAAGGCCACCTGCTCAACATCTTGAAGCCCTTATTTACTTGCTGTAAGAACAC 1828
 QY 581 AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrrAlaAspGlnSerIle 600
 Db 1829 AACCAAGATTTCTTTTGGGATGAGTACCCACATGAGATGCAATATGCAACAAAGCATC 1888
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrrGluTrpAsnAspAsn 620
 Db 1889 AAAGTGAAGATTAACCTTAATATGCTCTTGGAGATTAAGCATTAATGATGATGATGATGAT 1948
 QY 621 GluMetTyrrLeuPheArgSerSerValAlaTyrrAlaMetArgGlnTyrrPheLeuLysVal 640
 Db 1949 GAAATGTACTCTGTCGATCATCTGTTGCAATGCTATGAGGAGCATCTTTTAAAAAGTA 2008
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
 Db 2009 AAAATATGATGATTTCTTTTGGGAGAGATGTGCGATGCTATTTGAAACCAAGA 2068
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680


```

Db      2069 AATCTCTTAATTTCTTCTTCTACCTGCACCTAAAAATGCTGTGATATCATTTCTTGAACCT 2128
Qy      681 GIUVALGULYSALALLEARMETSERARGILEASNPALAPHEARGLEUASN 700
Db      2129 GAAGTTGAAGAGCCATCAGATGTCCTCCGACCCGATCAATGATGCTTCCGTGAAAT 2188
Qy      701 ASPASNSEULEUGLUPHELEUGLYILEGINPROTHLEUGLYPROPROASNGLPROPRO 720
Db      2189 GACACACAGCCTAGATGTTCTGGGAGATACAGCCACACTTGACCTTCAACAGCCCTCT 2248

RESULT 2
US-10-158-847-141
; Sequence 141, Application US/10158847
; Publication No. US20030091557A1
; GENERAL INFORMATION:
; APPLICANT: Tom Parry et al.
; TITLE OF INVENTION: Method and Compositions for Modulating ACE-2 Activity
; FILE REFERENCE: PP557
; CURRENT APPLICATION NUMBER: US/10/158,847
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 60/295,004
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 158
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 141
; LENGTH: 3396
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-158-847-141

Alignment Scores:
Pred. No.: 0 Length: 3396
Score: 3869.00 Matches: 720
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9 Indels: 0 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x US-10-158-847-141 (1-3396).
Qy      1 SerThrIleGluGlnAlaIysThrPheLeuAspLysPheAsnHisGlnAlaIuAsp 20
Db      136 TCCACCTTAGAGAGACAGCCAGACATTTTGGACAAAGTTTAACCAAGAGCCGAAAGC 195
Qy      21 LeuPheTyrGlnSerSerLeuAlaSerTyrAsnThrAsnIleThrGluGluAsn 40
Db      196 CTGTTCTATCAAAAGTTCCTGCTTCTGGAATTAATACCAATATTAATTAAGAGAAAT 255
Qy      41 ValGlnAsnMetAsnAlaIysAspLysThrSerAlaPheLeuLysGluGlnSerThr 60
Db      256 GTCCAAACATGAATTAATGCTGGGACAAATGGTCTCTTTTAAAGAAACAGTCCACA 315
Qy      61 LeuAlaGlnMetTyrProLeuGlnIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
Db      316 CTTCGCCAAATGATTCACATCAAGAAATTCAGAAATTCACAGTCAAGCTTCAGCTCAG 375
Qy      81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
Db      376 GCTCTCAGCAAAATGGGCTTTCAGTCTCTCAGAAACAAAGCAAAACGGTTGAAACA 435
Qy      101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAspPro 120
Db      436 ATTCTAAATACAAATGAGACCAATCTACAGTACTGAAAGTTTGTAAACCCAGATTAATCCA 495
Qy      121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
Db      496 CAAAGATAGCTTATTACTTGAACACAGGTTTGAATTAATGCAAAACAGTTTGAAGTAC 555
Qy      141 AsnGluArgLeuTyrPalaIleTyrPalaIleTyrPalaIleTyrPalaIleTyrPalaIle 160
Db      556 AATAGAGGGCTCTGGCTGGGAAAGCTGAGATCTAGGTGGGCAAGCAGCTGAGGCCA 615
Qy      161 LeuTyrGlnGluTyrValIleValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180

```

```

Db      616 TTATATGAGAGATATGAGTCTTAAAAATAGATGGCAGAGCAATCATATGAGAGC 675
Qy      181 TyrGluAspTyrTyrPalaIleTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
Db      676 TATGGGATTTATGGAGAGAGACTATGAACTAATGGGAGTATGATGCTATGACTACAGC 735
Qy      201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
Db      736 CGCGGCGAGTGTATGAAGATGAGATGACATCTTTGAAAGATTAACCATTTATATGAA 795
Qy      221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
Db      796 CATCTTCAATGCCATATGAGGCAAAAGTTGATGAATGCTTCTCTCATATACAGTCCA 855
Qy      241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheTyrPheAsnLeu 260
Db      856 ATTGATGCTCCTCCCTGCTCATTTGGTGTGATATGTTGGGTATGATTTGGACAAATCTG 915
Qy      261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
Db      916 TACTCTTGAACAGTCCCTTTGGACAGAAACCAATAGATGTTACTGATGATGATGGTG 975
Qy      281 AspGlnAlaTyrPalaIleArgIlePheLysGluAlaGluLysPhePheValSerVal 300
Db      976 GACCAAGCCTGGGATGCACAGAAATATTCAGAGAGCCGAGAGTTCTTGTATCTGTT 1035
Qy      301 GlyLeuProAsnMetThrGlnGlyPheTyrPalaIleAsnSerMetLeuThrAspProGlyAsn 320
Db      1036 GGTCTTCTCAATATATGATCAAGATTCAGGAAATTCATGCTAACGAGCCAGGAAAT 1095
Qy      321 ValGlnLysAlaValCysHisProThrAlaTyrPalaIleGlyLysGlyAspPheArgIle 340
Db      1096 GTTTCAGAAAGCGTCTGCCATCCACAGCTGGGAGCTGGGAGAGGCGACTTCAGATC 1155
Qy      341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHis 360
Db      1156 CTATGTGCAAAAGTGCATATGACAGCACTCTGACAGCTCATCATGATGGGGCAT 1215
Qy      361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
Db      1216 ATCCAGTATGATATGATGATGATGCTGACAACTTTCTGCTAGAAAGAGAGCTATATA 1275
Qy      381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProHisHisLeu 400
Db      1276 GGAATTCATGAAAGCTGTTGGGAAATCATGTACTTTCGCGACCCACACTAGCATTTA 1335
Qy      401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
Db      1336 AATTCATTTGGTCTTCTGTCACCCGATTTTCAAGAAACAAATGAAGAAATTAACCTTC 1395
Qy      421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuLys 440
Db      1396 CTGCTCAAAACAGCACTCAGATGTTGGGAGCTGCGCAATTAATCAATGTTAAGAGAG 1455
Qy      441 TyrArgTyrMetValPheLysGlyGluIleProLysAspGlnTyrPheLysTyrTyrP 460
Db      1456 TGGAGGTGATGCTCTTTAAAGGGAATTTCCCAAGACATGATGATGAAAGAGTGGTG 1515
Qy      461 GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCys 480
Db      1516 GAGATGAGAGAGATAGTGGGAGTGGGAGAACTGTGCCCATGATGAACATACAGTGT 1575
Qy      481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArg 500
Db      1576 GACCCGCGATCTCTGTTCAATGTTTAATGATCTCAATCAATGATTAATACACAGAG 1635
Qy      501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
Db      1636 ACCCTTACCAATTCAGATTCAGAAAGCACTTGTCAAGAGCACTTAACATGAGGCCCT 1695
Qy      521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540

```

```

US-09-978-385-2_COPY_19_738 (1-720) x US-10-114-893-85 (1-3325)

Best Local Similarity: 99.86% Mismatches: 0
Query Match: 99.90% Indels: 0
DB: 9 Gaps: 0

```

QY	1	SetThrIleGluIugInAlaIysThrPheLeuAspIysPheAsnHISGluIaGluAsp	20
Db	73	TCACCACTTAGGAGACAGGCCAAGACATTTTGGACAAAGTTTAAACCAAGACCGAAGAC	133
QY	21	LeuPheIyrInSerSerLeuAlaSerTrpAsnIyrAsnThrAsnIleThrIugIuAsn	40
Db	133	CTGTTCATCAAAAGTTCACCTTGCTTCCTGGAAATTAATTAACCAATATTACTGAAGAGAT	192
QY	41	ValGlnAsnMetAsnAsnAlaGlyAspIyrTrpSerIaphLeuIysGluInSerThr	60
Db	193	GTCCAAAACATGATAATATGCTGGGACAAATGCTTGCCTTTTAAAGNAAGTCCACA	252
QY	61	LeuAlaGlnMetIyrProLeuGlnIuIleGlnAsnLeuThrValIysLeuGlnLeuGln	80
Db	253	CTTGCCCAATGATATCCACTACACAAAGAAATTCAGAAATCTCACTCAAGCTTCAGCTGCAG	312
QY	81	AlaLeuGlnGlnAsnGlySerSerValIleSerGluAspIyrSerIyrAsnIleThr	100
Db	313	GCCTTCACCAAAATGGCTTCTAGTGCCTCCGAAAGACAAAGCAACGGTTGAACACA	372
QY	101	IleLeuAsnThrMetSerThrIleIyrSerThrGlyIysValCysAsnProAspAsnPro	120
Db	373	ATTCTAAATACATATGAGCACACATCTACAGTACGTGAAGAAAGTTGTAAACCAAGATATCCA	432
QY	121	GlnGluCysLeuLeuLeuGluIupProGlyLeuAsnGluIleMetAlaAsnSerLeuAspIyr	140
Db	433	CAAGATGCTTATTACTTGAACCAAGTTTGAATGAATATATGGCAAAAGTTTAGCTAC	492
QY	141	AsnGluIyrLeuTrpAlaTrpGluSerTrpArgSerGluValGlyGlnLeuAlaPro	160
Db	493	AATGAGAGCTCTGGGCTTGGCAAAACCTGCAGATCTGAGGTGGCAAGCAAGCTGAGGCCA	552
QY	161	LeuIyrGlnGluIyrValIleLeuIysAsnGluMetAlaIArgAlaAsnHISIyrGluAsp	180
Db	553	TATATGAGAGATATAGTGTCTTGAAATATGAGATGGCAAGCAAAATATTATGAGGAC	612
QY	181	IyrGlyAspIyrTrpArgIylAspIyrGluValAsnGlyValAspGlyIyrAspIyrSer	200
Db	613	TATGGGGATTATTGAGAGAGACATATGAATGAATGGGGTATGAGCTATACATACAC	672
QY	201	ArgGlyGlnIleIleGluAspValGlnHISThrPheGluGluIleIysProLeuIyrGlu	220
Db	673	CGGGCCAGTATGAGAGATGTGGAACTCTTGGAGAGATTAACCAATTATATGAA	732
QY	221	HISLeuHISAlaIyrValArgAlaIysLeuMetAsnAlaIyrProSerIyrIleSerPro	240
Db	733	CATCTTCATGCCTATGTAGGAGGCAAAAGTTGATGAATGCTATCCTCTATATACGTCA	792
QY	241	IleGlyCysLeuProAlaHISLeuLeuIylAspMetTrpGlyIyrArgPheTrpThrAsnLeu	260
Db	793	ATTGGATGCTCTCCCTGCTCATATTTGCTTGATATGTGGAGATTTTGGCAAAATCTG	852
QY	261	TyrSerLeuThrValProPheGlyGlnIysProAsnIleAspValThrAspAlaMetVal	280
Db	853	TACTCTTGACAGTCCCTTGGACAGAAACCAACATATGATGTTACATGATCCAAATGCTG	912
QY	281	AspGlnAlaTrpAspAlaGlnArgIlePheIylsGluAlaGluIysPhePheValSerVal	300
Db	913	GACCAAGGCTGGAGATGCACAGATATATTCAGAGGCGGAGAAATCTTGATATCTGT	972
QY	301	GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn	320
Db	973	GGTCTCTCTATATGACACCAAGATCTCGGAAAAATTCATAGCTTACAGGACCCAGGAAT	1032
QY	321	ValGlnIysAlaValCysHISProThrAlaTrpAspLeuGlyIysGlyAspPheArgIle	340
Db	1033	GTTTAGAAGACAGTCTGCATCCCAAGCTTGGACCTGGGGAAGGCGGATTCAGAGATC	1092

OY	341	LeuMetCysThrIysValIleThrMetAspSerPheLeuThrIaHisSHISgluMetGlyHis	360
Db	1093	CTTAATGGCACAAAGGTGACATGGAGACTCTGTGACAGCTCATCGAGATGGGGCAT	1152
OY	361	IleGlnIYrAspMetIaIaYrIaIaGlnProPheLeuLeuYrAsnGlyIaIaAsnGlu	380
Db	1153	ATCCAGATATGATATGGCATATGGTGCACAACCTTTCTGGCTAAAGAAATGGAGCTAATGAA	1212
OY	381	GlyPheHisGluIaIaValGlyGluIleMetSerLeuSerIaIaIaThrProIyHisIleu	400
Db	1213	GGATTCATGATGAGGTGGTGGGAAATATGCACTTTCTCGAGCCACACCTTAACCATTTA	1272
OY	401	LysSerIleGlyLeuLeuSerProAspPheGlnIuAspAsnGlnIYrGluIleAsnPe	420
Db	1273	AAATCCATGGTCTTCTGTCTACCCGATTTTCAAAAGACATGAAGAAATAACTTC	1332
OY	421	LeuLeuIysGlnIaIaLeuThrIleValGlyThrLeuProPheThrIYrMetLeuGluys	440
Db	1333	CTGCTCAAAACAGCACACACCATGTTGGGACTGTGCATTTACTTACATAGTTTGAAGAG	1392
OY	441	TrpArgTrpMetValaPheIysGlyGluIleProIyAspGlnIYrPmetIySyrTrp	460
Db	1393	TGGAGGTGATGGTCTTTAAAGGGAAATTTCCAAAGACAGTGGATGAAGAAATGGTGG	1452
OY	461	GluMetIysArgGlnIleValaGlyIaIaIuProValProHisAspGluThrIYrCys	480
Db	1453	GAGATGAAGCGAGATAGTGTGGGTGGTGAACCTGGCCCATGATGAACATACGT	1512
OY	481	AspProAlaSerLeuPheHisValSerAsnAspIYrSerPheIleIaGlyTYrThrArg	500
Db	1513	GACCCCGCATCTGTTCCATGTTTCAATATTAATCAATTCATTCATTCATTAATTAACAAG	1572
OY	501	ThrLeuTYrGlnPheGlnPheGlnIaIaLeuGysGlnIaIaIaIyHisGluIPro	520
Db	1573	ACCCTTACCAATTCACATTCACAAAGACATTTGTCAAGACGATAAACATGAAGGCCCT	1632
OY	521	LeuHisIyCysAspIleSerAsnSerThiGluIaGlyGlnIySleuPheAsnMetLeu	540
Db	1633	CTGCACAATGTGACATCTCAACACTCTACACACCTGGACAGAAACGTTCMAATATGCTG	1692
OY	541	ArgLeuGlyIySerGluProTrpThrLeuIaIeugIuAsnValIaIaGlyIaIaIyAsn	560
Db	1693	AGCGTTGAAATTCAGAACCTCGACCTTACCATTTGGAAATGTGTTAGAGGACAAAGAC	1752
OY	561	MetAsnValaIArgProLeuLeuAsnTYrPheGluProLeuPheThrTriPLeuIyAspGln	580
Db	1753	ATGATATGTAGGCCACCTGCACACTTGTAGGCCCTAATTACTGTGGCTGAAGACAG	1812
OY	581	AsnIyAsnSerPheValaGlyTrpSerThrAspTrpSerProTYrIaIaAsnGlnSerIle	600
Db	1813	AACAAGAAATCTTTGTTGGATGGATGACACACAGTGAATTCATATGACACCAAGACATC	1872
OY	601	LysValaIArgIleSerLeuIySerIaIaIeugIyAspIySAlaTYrGluTrpAsnAspAsn	620
Db	1873	AAAGTGAAGATAGCGCTAAATACAGCTTTGGAGATTAAGCATATGATGAGACAGACAT	1932
OY	621	GluMetTYrLeuPheArgSerSerValaIaTYrIaIaMetArgGlnTYrPheLeuIyVala	640
Db	1933	GAATATGTACCTGTTCCGATCATCTGTTCCATATGCTATAGAGCGAGTACTTTTAAAGTA	1992
OY	641	LysAsnGlnMetIleLeuPheGlyGluGluAspValaIArgValaIaAsnLeuIyProArg	660
Db	1993	AAAAATCAGATGATCTTTTGGGAGAGAGCATGTGCCAGTGGCTAATTGAACCAAGA	2052
OY	661	IleSerPheAsnPhePheValaIThrIaProIyAsnValaIserAspIleIaProArgThr	680
Db	2053	ATCTCTTAAATTTCTTTGTCACTGCACCTAAATATGTCTGATATCATTTCTAGAACT	2112
OY	681	GluValaGlyIySAlaIleArgMetSerIySerArgIleAsnAspIaIaPheArgLeuAsn	700
Db	2113	GAAGTGAAGAAAGCCATCAGATGATGTCCGGAGCCGTATACATGATGCTTCCGTGTGAT	2172

```

QY      701  AspsasSerIengInphetengLylegInprothIengLylProPoaangInproPro  720
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      2173 GACAACGCGCTAGAGATTCTGGGGGATACAGCCACCACTTGGACCTCTTAACCAAGCCCT  223

RESULT 4
US-09-969-384-2
: Sequence 2, Application US/09969384
: Publication No. US20020192749A1
: GENERAL INFORMATION:
:   APPLICANT: Moore, et al.
:   TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
:   FILE REFERENCE: P055P1
:   CURRENT APPLICATION NUMBER: US/09/969,384
:   PRIOR FILING DATE: 2001-10-03
:   PRIOR APPLICATION NUMBER: PCT/US01/10542
:   PRIOR FILING DATE: 2001-04-02
:   PRIOR APPLICATION NUMBER: 60/236,384
:   PRIOR FILING DATE: 2000-09-29
:   PRIOR APPLICATION NUMBER: 60/194,118
:   PRIOR FILING DATE: 2000-04-03
:   NUMBER OF SEQ ID NOS: 27
:   SOFTWARE: PatentIn Ver. 2.0
:   SEQ ID NO 2
:   LENGTH: 2920
:   TYPE: DNA
:   ORGANISM: Homo sapiens
:   FEATURE:
:     NAME/KEY: SITE
:     LOCATION: (1707)
:     OTHER INFORMATION: n equals a,t,g, or c
:     NAME/KEY: SITE
:     LOCATION: (2702)
:     OTHER INFORMATION: n equals a,t,g, or c
:     NAME/KEY: SITE
:     LOCATION: (2749)
:     OTHER INFORMATION: n equals a,t,g, or c
:     NAME/KEY: SITE
:     LOCATION: (2757)
:     OTHER INFORMATION: n equals a,t,g, or c
:     NAME/KEY: SITE
:     LOCATION: (2788)
:     OTHER INFORMATION: n equals a,t,g, or c
:     NAME/KEY: SITE
:     LOCATION: (2789)
:     OTHER INFORMATION: n equals a,t,g, or c
:     NAME/KEY: SITE
:     LOCATION: (2819)
:     OTHER INFORMATION: n equals a,t,g, or c
:     NAME/KEY: SITE
:     LOCATION: (2835)
:     OTHER INFORMATION: n equals a,t,g, or c
:     NAME/KEY: SITE
:     LOCATION: (2856)
:     OTHER INFORMATION: n equals a,t,g, or c
US-09-969-384-2

Alignment Scores:
Pred. No.:      0      Length:      2920
Score:          3833.00  Matches:      715
Best Local Similarity: 99.31%   Conservative: 1
Query Match:      99.17%   Mismatches:  4
DB:              99.07%   Indels:      1
                       Gaps:      0

US-09-978-385-2_COPY_19_738 (1-720) x US-09-969-384-2 (1-2920)
QY      1  SerThrIlegIugIugInAlaIaIySIRPhleu-AsplysPheasnhISgluaIagIuAs  20
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      83  TCCACCATTCAGGAGAACGCGCCAGACATTTTGGGACAGTTTAAACCGAAGCCGAGAA  142

20  pIenPhetIcInserSerIeunIaserTPrpansTyrasnprranIlethngIugIuAs  40
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      143 CCGTGCTATCAAGATTCCTCTCTGGAAITTAACCAACCAATATTAAGTGAAGAGA  202

```

QY 40 nValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerTh 60
 |||||
 Db 203 TGTCCAAAACATGATGATGCTGGGACAAATGCTCTCTTTTAAAGGAAGCTCAC 262
 QY 60 rLeuAlaGlnMetLysTrpProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuG 80
 |||||
 Db 263 ACTTGCCCAAAATGATGCTGACATCAAGAAATTCAGAAATCTCAGAGCTTCAAGCTTCCA 322
 QY 80 nAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnTh 100
 |||||
 Db 323 GGCCTTCAGCAAAATGGGTCTTCAGTCTCTCAGAAACAGCAAGCAACCGTTGACAC 382
 QY 100 rIleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPr 120
 |||||
 Db 383 AATTCCTAAATACAAAGACACCATCTACAGTACTGAAAGTTTGTAAACCCAGATTAATCC 442
 QY 120 oGlnGlnCysLeuLeuLeuGlnLupProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTy 140
 |||||
 Db 443 ACAAGATGCTTATCTGTAACCCAGTTTGAATGAATGAATGAATGAATGAATGAATGA 502
 QY 140 rAsnGlnArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPr 160
 |||||
 Db 503 CAATGAGAGGCTGTGGGCTTGGGAAAGCTGAGATCTGAGTCCGCAAGCAAGCTGAGGCC 562
 QY 160 oLeuTyrGlnGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAs 180
 |||||
 Db 563 ATTATATGAGAGATGATGCTCTGCAAAATGAGATGAGCAAGCAAAATCAATATGAGGA 622
 QY 180 PTYrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSe 200
 |||||
 Db 623 CTATGGGATTTATGAGAGAGAGACTATGAATGAATGAATGAATGAATGAATGAATGA 682
 QY 200 rArgGlyGlnLeuIleGluAspValGlnHisThrPheGlnGluIleLysProLeuTyrG 220
 |||||
 Db 683 CCGCGGCGAGTGTGATGAGATGAGAACATCTTGAAGAGATTAACCATATATATGA 742
 QY 220 uHisLeuHisAlaTyrValAlaArgAlaLysLeuMetAlaLysTrpSerTyrIleSerPr 240
 |||||
 Db 743 ACATCTTATGCTTATGAG 802
 QY 240 oIleGlyCysLeuProAlaHisLeuLeuGlnLysAspMetTrpGlyArgPheTrpThrAsnLe 260
 |||||
 Db 803 AATGGATGCTCCCTGCTCATTTGCTGTGATATGATGGGGTATGATTTGGACAAATYT 862
 QY 260 uTyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVa 280
 |||||
 Db 863 GTACMSTTTGACAGTTCCTTGGACAGAAACCAACATAGATGTCTGATGCAATGGT 922
 QY 280 lAspGlnAlaTrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerVa 300
 |||||
 Db 923 GGACCAAGRCCCTGGATGACAGAGAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 982
 QY 300 lGlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAs 320
 |||||
 Db 983 TGTCTCTCTTAATATGACTCAAGATCTGGGAAATTCATGCTAAGGACCCAGAGAAA 1042
 QY 320 nValGlnLysAlaValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArg 340
 |||||
 Db 1043 TGTTCAGAAAGCACTGCTGCAATCCACAGCTTGGAGACTGGGGAGGGGCACTTCAGAT 1102
 QY 340 eLeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisIleGlnMetGlyH 360
 |||||
 Db 1103 CCTTATGTGCAGAAAGGTGACATGACGACTTCTCTACAGCTCATATAGATGGGGCA 1162
 QY 360 sIleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnG 380
 |||||
 Db 1163 TATCCAGATGATATGCAATATGCTGCAACACTTTTCTCTAAGAAATGAGACTATGA 1222
 QY 380 uGlyPheHisGlnAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLe 400
 |||||
 Db 1223 AGATTCATGATGAAGCTGTGGGAGAAATCATGCTTCTGACAGCCACACCTAAGCATTT 1282

QY 400 uLysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPh 420
 |||||
 Db 1283 AAAATCATTTGGTCTTGTGACCCGATTTTCAAGAGACAAATGAAGAAATTAACAT 1342
 QY 420 eLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGly 440
 |||||
 Db 1343 CTCGTCAAAACAGCACTGACAGATGTGGGACTGTGCCATTTACTTACATGTTAGAGA 1402
 QY 440 sTrpArgTrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTrpTr 460
 |||||
 Db 1403 GTGAGGTGGATGGTCTTTTAAAGGGGAAATTCCTCAAAAGCACTGATGATAAAAGTGTG 1462
 QY 460 pGluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCy 480
 |||||
 Db 1463 GGAGATGACGAGAGATGATGAGGTGGGAGAGCTGTGCCCATGATGAACATACTG 1522
 QY 480 sAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrAr 500
 |||||
 Db 1523 TGACCCCGCACTCTCTCTCCATGTTCTTAATGATTAATCATTCATTCATTCATTCACAG 1582
 QY 500 gThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGlyLys 520
 |||||
 Db 1583 GACCCCTTACCAATTCAGATTCAGAGAGCACTTTGTCAAGCAAGCTTAAACATGAAGGCC 1642
 QY 520 oLeuHisLysCysAspIleSerAsnSerThrGluAlaGlnLysLysLeuPheAsnMetLe 540
 |||||
 Db 1643 TCTGCAAAATGAGATCTCAAACTCTCAAGAGAGTGGAGCAAGAACTTCAATATGCT 1702
 QY 540 uArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGlnLysAsnValAlaGlyAlaLysAs 560
 |||||
 Db 1703 GAGGNTTGAAGAAATCAAGAACCCCTGACCTTACATGGAATGTTGTAGGAGCAAGAA 1762
 QY 560 nMetAsnValArgProLeuLeuAsnTyrPheGlnLupProLeuPheTrpLeuLysAspG 580
 |||||
 Db 1763 CATGATGTAAGGCCACTGCTCAACACTTTGAGCCCTTATTTACTGCTTAAAGACCA 1822
 QY 580 nAsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSer 600
 |||||
 Db 1823 GAACAAAGAAATCTTTTGTGGAGTGAAGTACCGACCTGAGATGCAATGAGCAAAAGCAT 1882
 QY 600 eLysValArgIleSerLeuLysSerAlaLeuGlnLysAspLysAlaTyrGluTrpAsnAsp 620
 |||||
 Db 1883 CAAGTGAAGATTAAGCTTAATCAACCTTGTGAGATTAAGATTAAGATTAAGATTAAG 1942
 QY 620 nGluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVa 640
 |||||
 Db 1943 TGAATGTACCTGCTCCGATCATCTGTTGCATATGCTATGAGAGCACTTTTAAAGT 2002
 QY 640 lLysAsnGlnMetIleLeuPheGlyGluLysAspValArgValAlaAsnLeuLysProAr 660
 |||||
 Db 2003 AAAAATTCAGATGATCTTTTGGGGAGAGAGATGAGAGAGCTTAATTTGAACCAAG 2062
 QY 660 gIleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgTh 680
 |||||
 Db 2063 AATCTCTTAAATTTCTTGTCTACCTGCACTTAAATATGCTGTATATCATCTTCAGAAC 2122
 QY 680 rGluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAs 700
 |||||
 Db 2123 TGAAGTTGAAGAGCATCAGATGCGCGAGCGCTATCATATGCTTCCGCTGAA 2182
 QY 700 nAspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPr 720
 |||||
 Db 2183 TGAAGCAAGCTAGATGTTCTGGGATACAGCAACACTTGGACCTCTAAGCAGGCCCC 2242
 QY 720 o 720
 |||||
 Db 2243 T 2243

RESULT 5
 US-10-158-847-137
 ; Sequence 137, Application US/10158847
 ; Publication No. US20030091557A1
 ; GENERAL INFORMATION:

```

APPLICANT: Tom Parry et al.
TITLE OF INVENTION: Method and Compositions for Modulating ACE-2 Activity
FILE REFERENCE: PF557
CURRENT APPLICATION NUMBER: US/10/158,847
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: 60/295,004
PRIOR FILING DATE: 2001-06-04
NUMBER OF SEQ ID NOS: 158
SOFTWARE: PatentIn version 3.1
SEQ ID NO 137
LENGTH: 2920
TYPE: DNA
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1707)..(1707)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2702)..(2702)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2749)..(2749)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2757)..(2757)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2788)..(2788)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2819)..(2819)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2835)..(2835)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2856)..(2856)
OTHER INFORMATION: n equals any amino acid
US-10-158-847-137
Alignment Scores:
Pred. No.: 0
Score: 3833.00
Percent Similarity: 99.31%
Best Local Similarity: 99.17%
Query Match: 99.07%
Length: 2920
Matches: 715
Conservative: 1
Mismatch: 4
Indels: 1
Gaps: 0
US-09-978-385-2_COPY_19_738 (1-720) x US-10-158-847-137 (1-2920)
QY 1 SerThrIleGluGlnAlaIysThrPheLeu-AspLysPheAsnHisGluAlaGluAs 20
Db 83 TCACACATTGAGGAAAGCCAGACATTTTGGACAAAGTTTAACACAGAACCCGAGAA 142
QY 20 pLeuPheTyGlnSerSerLeuLaserThrPasnTyrrAsnThrAsnIleThrGluLys 40
Db 143 CCGTCTTATCAAAAGTTCACCTTCCTTGGAAATTTAAACACCAATATTAAGTAAAGAA 202
QY 40 nValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSer 60
Db 203 TGTCCAAAACATATAATATCTGCGGACAAATGCTGCTTTTAAAGAACAGTCCAC 262
QY 60 rLeuAlaGlnMetTyrrProLeuGlnIleGlnAsnLeuThrValLysLeuGlnLeuG 80
Db 263 ACTTGCCCAAAATGTATCCACTACAGAAATTCAGAAATTCACAGTCAAGCTTCACTGCA 322

```

```

QY 80 nAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsn 100
Db 323 GGCCTTACGAAATAATGGGCTCTCAGTCTCCAGAAACAGAGCAACGGTTGAACAC 382
QY 100 rLeuAsnThrMetSerThrIleTySerThrGlyValAlaLysAsnProAspAsn 120
Db 383 AATTCTAATACAAATGACACCATCTACAGTACTGGAAAAATTGTAAACCAATATCC 442
QY 120 oGlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAsp 140
Db 443 ACAAGAAATGCTATTACTTGAACACAGTTTGAATGAATATATGCAAAACAGTTAGACTA 502
QY 140 rAsnGluArgLeuThrAlaTrpGluSerTrpArgSerGluValGlySerGlnLeuAsp 160
Db 503 CAATGAGAGCTCTGGGCTGGGAAAGCTGAGATCTGAGTCCGCAAGCAGCAGAGCC 562
QY 160 oLeuTyrgluGluTyrrValIleLeuLysAsnGluMetAlaArgAlaAsnHisTyrglu 180
Db 563 ATTATATGAAAGTATGTGCTTGAATAAATGAGATGGCAAGCAAAATCTTTTGAAGA 622
QY 180 pTyrgluAspTyrrTrpArgGlyAspTyrgluValAsnGlyValAspGlyTyrrAsp 200
Db 623 CTATGGGATATATGAGAGAGAGACTATGAAGTAAATGGGTAGATGCTATGACTACAG 682
QY 200 rArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrg 220
Db 683 CCGCGCCAGTTGATGGAAGATGGGAACATTCCTTGAAGAGATTAACCATATATGA 742
QY 220 uHisLeuHisAlaTyrrValArgAlaLysLeuMetAsnAlaTyrrProSerTyrrIle 240
Db 743 ACATCTTCATGCTATGTGAGGCCAAATGTATGATCCCTTCTCTATACAGTCC 802
QY 240 oIleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheThrAsnLe 260
Db 803 AATTGAGATGCCCTCCCTGCTCATTTGCTGGATATGAGGATGATTTGGACAAATYT 862
QY 260 uTyrrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMe 280
Db 863 GTACWSTTTGACAGTTCCTTTGGACAGAAACCAACATGATGTATGATGGAATGCT 922
QY 280 lAspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaLysPhePheAlaSerVa 300
Db 923 GGACCAAGRCCTGGATGCAACAGAAATTTCAAGAGGCCAGAAATTTCTTGTATCTGT 982
QY 300 lGlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGly 320
Db 983 TGGTCTTCCTAATATGACTCAAGAGATCTGGGAAATTCATGCTAACGAGCCAGGAA 1042
QY 320 nValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheAla 340
Db 1043 TGTTCAGAAAGCATCTGCCATCCCAAGCTTGGACCTGGGGAAGGCGCACTTCAGAT 1102
QY 340 eLeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGly 360
Db 1103 CCTTATGTGCACAAAGGTGACATGTGAGAGACTCTCTGACAGCTCATGTAGATGGGCA 1162
QY 360 sIleGlnTyrrAspMetAlaTyrrAlaAlaGlnProPheLeuLeuArgAsnGlyAla 380
Db 1163 TATCCAGATATGATATGCAATATGCTGCACAACTTTTCTGTAAGAAATGAGCAATGA 1222
QY 380 uGlyPheHisGluAlaValGlyGluIleMetSerLeuSerLeuAlaThrProLysHis 400
Db 1223 AGGATTCATGAAAGCTGTGGGGAATATGATGCTATTCCTGACGACCACTTAAGCA 1282
QY 400 uLysSerIleGlyLeuLeuSerProAspPheGlnLysAsnGluThrGluIleAsnPh 420
Db 1283 AAAATCCATGTGCTTCTGTCACCGATTTTCAAAAGCAATGAAGAAATTAATTT 1342
QY 420 eLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrrMetLeuGly 440
Db 1343 CCTGCTCAAAACAGACACTCAGATTTGTTGGACTCTGCCATTTACTTACATGTTAGA 1402
QY 440 sTrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysTrpTr 460

```

```

Db      1403 GTGGAGGCTGATGCTTTAAAGGGGAAATTCACCAAGCCGATGAGAAAGTGGTG 1462
Qy      460 pglumetlysargluilevalglvalvalgluprovalprohisaspgluthrtycy 480
Db      1463 GGAGATGAAGCAGAGATAGTGGGGTGGTGAACCTTGCCCATGTATGAACATATGTG 1522
Qy      480 saspaprolasertleuphehisvalserasnasptryserphellearglytyrthar 500
Db      1523 TGACCCCGCATCTGTTCCAGTTCTTAAGATATACATTCATTCATTCATTCATTCACAA 1582
Qy      500 gthleutyrglnpheglnpheglnlualaleucysglnlaalaalysghsluglypr 520
Db      1583 GACCCCTTACCAATTCAGTTCCAGAAAGCACTTTGTCTAAGCAGCTAAACATGAAGGCC 1642
Qy      520 oleuhislyscysapplieserasnserthrghualagllyglnlyseupheasnmetle 540
Db      1643 TCTGCACAAATGTGACATCTCAACTACAGAAAGCTGCAGAAACGTTCCATATTCCT 1702
Qy      540 uargleuglylyserserluprotptthrleualaleugluasnvalvalglvalalysas 560
Db      1703 GAGGWTGGAAATCAGAACCCCTGCACCTAGCATGGAAATGTTGTAGAGCAAGAA 1762
Qy      560 nmetasnvalargproleuleuantryrphelglnproleupherthrleupheasnasp 580
Db      1763 CATGAATGTAGGCCCTGCTCACTCACTTGAAGCCCTTATTCCTGGCTGAAAGACCA 1822
Qy      580 nasulysasnserphevalgltyrpsertthrastpserprotyrilaaspglinser1 600
Db      1823 GAAACAAGAAATCTTTTGGGATGGAGTACGACGAGTGCATATGACAGACCAAGCAT 1882
Qy      600 elysvalarglieserleuylseralaleuglyaspysalatyrgluthrpaasnaspas 620
Db      1883 CAATGTGAGGATTAACCTTAATCAGCTCTTGAGATTAACCATATGATGAACGACCA 1942
Qy      620 nglumetlyrleupheargsersevalalatyralametrarglntryrphelleyva 640
Db      1943 TGAATGTACTCTTCCCATATCTGTGCTATAGCTATAGGAGCATCTTTTAAAGT 2002
Qy      640 llyasnglmethylleuphegilygluaspvalargvalalasnleuylsproar 660
Db      2003 AAAAATCAGATGATCTTTTGGGAGAGAGATGGAGTGGCTAATTTGAANACAG 2062
Qy      660 gtlseerpheasnpherphevalthrallaprollysasnvalseraspilleproagth 680
Db      2063 AATCTCTTAATTTCTTGTGCTACTGCACCTAAATGTCTGATATCATTTCTTAAGAC 2122
Qy      680 rgluvalglutysalalalearmetserargserargileasnaspilapheargleuas 700
Db      2123 TGAAGTTAAAAAGCCATCAGAGATGTCGCCGAGCCGATTCATATGCTTTCCGTCGAA 2182
Qy      700 naspasnserleuglnpheuglylleglnprothrleuglyproproasnnglnpropr 720
Db      2183 TGACGACAGCCTAGAGTTCTTGCGGATACAGCAACACTTGAGACTCTCTAACGAGCCCC 2242
Qy      720 o 720
Db      2243 T 2243

```

```

RESULT 6
US-09-969-384-12
; Sequence 12, Application US/09969384
; Publication No. US20020192749A1
; GENERAL INFORMATION:
; APPLICANT: MOORE, et al.
; TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PRO55P1
; CURRENT APPLICATION NUMBER: US/09/969,384
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: PCY/US01/10542
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/2236,384
; PRIOR FILING DATE: 2000-09-25

```

```

; PRIOR APPLICATION NUMBER: 60/194,118
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 2911
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1705)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2847)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-969-384-12

Alignment Scores:
Pred. No.: 0
Score: 3785.00
Percent Similarity: 99.31%
Best Local Similarity: 99.17%
Query Match: 97.83%
DB: 9
Matches: 2911
Conservative: 1
Mismatches: 4
Indels: 3
Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x US-09-969-384-12 (1-2911)

Qy      1 SerThrIleGluGlnAlaLysThrPheLeu-AspLysPheAsnHisGluAlaGlns 20
Db      83 TCACACCTTAGAGACAGACGCCCAACATTTTGGACAAAGTTTAACCAAGCCGAAACA 142
Qy      20 PleuPheTyrglnSerSerleuAlaSerTrpAsnTyraAsnThrAsnIleThrGluGlns 40
Db      143 CCGTCTCATCAAAAGTTCACTTCTCTTGGAATTTATACCAACCAATATTACTGAAGACA 202
Qy      40 nValGlnAsnMetCAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerth 60
Db      203 TGTCCAAACATATAAATATGCTGGGACAAATGCTGCTTTTAAAGAACAGTCCAC 262
Qy      60 rLueAlaGlnMetTyrrProleuGlnIleGlnAsnLeuThrValLysLeuGlnLeuG 80
Db      263 ACTTGCCCAATATATCCACTACACAAATTTGCAATTCACAGTCACTTACCTGACGCA 322
Qy      80 nAlaLeuGlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnth 100
Db      323 GGCTCTTACCAAAATAGGGCTTCAGTGTCTTCAGAAACAGAGCAACGAGTTGACAC 382
Qy      100 rIleLeuAsnThrMetSerThrIleTyrrSerThrGlyLysValCysAsnProAspAsnPr 120
Db      383 AATTCCTAAATACATAGACACCAATCTACAGTACGTGAAAAGTTTGTAAACCCAGATATCC 442
Qy      120 OGlnGluCysLeuLeuLeuGlnIupProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTy 140
Db      443 ACAAGAAATGTTATTACTTGAAACCAAGTTTGAATGAATGAATGAATGAATGAATGA 502
Qy      140 rAsnGluArgLeuTrpAlaTrpGlnuSerTrpArgSerGlnValAlgllysglnLeuArgPr 160
Db      503 CAATGACAGGCTCTGGGCTTGGGAAGCTGAGATCTGAGTCCGGAACACAGCTAGAGGC 562
Qy      160 oleuTyrgluGlnTyrrValValleuLysasnGlnMetAlaArgAlaAsnHisTyrgLue 180
Db      563 ATTATATGAAGATATGCTGTTGAAAATGATGATGAGCAAGCAAAATCATATATAGAGA 622
Qy      180 pTyrgLysAspTyrrTrpArgGlyAspTyrglnValAlasnGlnValAspGlyTyrrAspTySe 200
Db      623 CTATGGGAGTATTATGGAGAGGAGACTATAGATGAATGAGGATGCTATGACTAAG 682
Qy      200 rArgGlyGlnLeuLeuGlnAspValAluHisThrPheGlnGlnIleLysProLeuTyrgl 220
Db      683 CCGCGGCGCAGTTGATTTGAAGATGTGGAACATACCTTTGAAGAGATTTAAACCTTATAGA 742
Qy      220 uHisLeuHisAlaTyrrValAlaArgAlaLysLeuMetAsnAlaTyrrProSerTyrrIleSerPr 240

```


Db 743 ACATCTTCATGCTGATGAGGCCAAGTGTGATGATGCTATCTTCTATATCAAGTCC 802
Qy 240 oilegylcysleuprolahisleuleuglyaspmetrpglyarphetrphrasnle 260
Db 803 AATTGGATGCTCCCTCTCATTTGCTGGTATATGTGGGTAGATTTTGGCAAAATCT 862
Qy 260 uyrserleuthvalprophelglylinsproasnleaspylthraspalametva 280
Db 863 GTACTCTTGTACATCTCTTGTGACAGAAACCAACNATGATGTCTATCTCATATGCT 922
Qy 280 laspghlatripspialaglnarqilephelysglnalaglylusphepheyalserva 300
Db 923 GGACAGAGCCCTGGATGACAGAGATATTCAGAGAGCCGAGAGATTTCTT-CTATCTGT 981
Qy 300 lglleuproasnmetrthglnlyphetrpgluasnsermetleuthrasploglyas 320
Db 982 TGGTCTTCTTAATGTGATCAAGATCTGGGAAAATTCATGCTAACGAGCCAGGAAA 1041
Qy 320 nvalglnlyalaalvalcysshsiprothrilaatpaspheuglylysglyaspheargyl 340
Db 1042 TGTTCAGAAAGCACTGTGCAATCCAGCTGGAGCTGGAGGAGGCACTTCAGAT 1101
Qy 340 eleumetcystrhlystvalthrmelaspaspheleuthralahislsiglumetglyl 360
Db 1102 CTTATGTGCAAAAGGTGACAAATGACGACTTCTTCTGACATCATATGATGGGCA 1161
Qy 360 silglnlyraspmetlalatryalaalaglnprophleuleuargasnlyalaasnl 380
Db 1162 TATCCAGATATGATGCAATGTGTCACAACTTTCTGCTAGAAATGAGAGCTATGA 1221
Qy 380 ugllypnehlsiglualvalglylilemetserleuserlialahthrprolyshlsle 400
Db 1222 AGGATTCATGATACCTGTGGGAAACATGCTACTTGTGACACCCACCTTAAGCATTT 1281
Qy 400 ulysserlileglyleuuserproaspheglnlaspasnlglnthrghluleasnph 420
Db 1282 AAAATCCATGTGCTTGTGTCACCCGATTTTCAAGAAAGACATGAACAAATTAACCTT 1341
Qy 420 eleuleulysglnalaleuthrlvalaglythrleuprophetrhrymetleugly 440
Db 1342 CCTGCTCAAAACACACACTGATGTTGGGACTGCTGCTCATTTACTTACATGTTAGAA 1401
Qy 440 strpargtrpmetvalahelysglyluleprolyaspghlntrpmetlystprtr 460
Db 1402 GTGGAGGTGATGCTTTTAAAGGGGAAATCCCAAGAGCAGGATGAAAGGCTGTG 1461
Qy 460 pglumetlysragslulevalaglylvalaglylprovalprohlsaspghlthrtrcy 480
Db 1462 GGAATGAGAGGAGATAGTGGGGTGGTGGAACTGTGCCCATGATGAACATATCTG 1521
Qy 480 saspprolaserleuphehlsvalaserasnspyrserphelleargytrrhar 500
Db 1522 TGACCCCGCATCTGTCTCCATGTTCTAAATGATTACATTCATGATGTATACACAG 1581
Qy 500 gthrleuthrlnpheglnpheglnlualaleucysghlalaalysghlsuglypr 520
Db 1582 GACCTTTTACCAATTCAGTTTCAAGAGCACTTGTCAACACACTTAACTGAAGGCC 1641
Qy 520 oleuhlslyscysaspliserasnserthglnalaglylinsleupheasnmetle 540
Db 1642 TCTGCAAAATGTGACATCTC-MACTCTACAGAAAGCTGGACAGAAAGCTTCAATATGCT 1700
Qy 540 uargleuglyllyssergluprotprthrleualaleuglnuasnvalaglylalyas 560
Db 1701 GAGGTTTGAAGAAATCAGAACCTGAGCCCTGACATGTGAAATGTGTGAGAGCAAGAA 1760
Qy 560 nmetasnvalargproleuasntrypheglnproleuphehtrtrpleuysasppl 580
Db 1761 CATGAATGTAAGGCCACGCTCACTACTTGTGACCCCTTATTTACCTGGCGAAGACCA 1820
Qy 580 nasnlyasnserphevalaglytrpserthrassprrpserprotyrilaaspghlser 600
Db 1821 GAACAAGATTTCTTTGTGGGATGAGTACCGACTGGAGTGCATATGACAGCAAGCAT 1880

Qy 600 elysvalarglleserleuylserseralaleuglyaspysalatyrglutrpsasnspas 620
Db 1881 CAAAGTAGAGATTAAGCCTTAATCAAGCTCTTGGAGATTAAGCATATGATGAGACGACAA 1940
Qy 620 nglumetlyrleupheargsersevalalalatrvalahetarglnlyrpheluleysva 640
Db 1941 TGAATGTGCTGTTCCTGATCAATCTGTGATATGATGAGGAGTACTTTTAAAGT 2000
Qy 640 llysaenglnmetlileupheglnlylulaspvalargvalalaasnleulysproar 660
Db 2001 AAAAATTCAGATGATTTCTTTGGGAGAGAGATGGCAGAGGCTTAATTTAAACCAAG 2060
Qy 660 gllserpheasnphelvalthrilaalprolyasnvalseraspilleleproargth 680
Db 2061 AATCTCTTAAATTTCTTGTGCTACGCTCACTAAATATGCTGATATCATCTTAAAGAC 2120
Qy 680 rgluvalaglylvalalelsermetsersearqileasnspalapharileus 700
Db 2121 TGAAGTTGAAGGCAATGATGATGTCGAGAGCCGATCAATGATGCTTCCGCTGGA 2180
Qy 700 naspasnerleuglupheleuglylileglnprothrleuglyproproasnlnpropr 720
Db 2181 TGACGACAGCCTTAGATTTCTGGGATTAAGCAACACTTGGACCTTAACACAGCCCCC 2240
Qy 720 o 720
Db 2241 T 2241

RESULT 7
US-10-028-072-71
Sequence 71, Application US/10028072
Publication No. US20030004311A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OF INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028, 072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19

PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910

PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090538
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07

 Alignment Scores:
 Pred. No.: 0 Length: 3732
 Score: 3720.00 Matches: 717
 Percent Similarity: 83.88% Conservative: 1
 Best Local Similarity: 83.76% Mismatches: 2
 Query Match: 96.15% Indels: 138
 Gaps: 1

 US-09-978-385-2_COPY_19_738 (1-720) x US-10-028-072-71 (1-3732)

QY 1 SerThrIleGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 94 TCACCATTTGAGGACAGCCAGACATTTTGGACAGATTACCCAGAGCCGAGAGC 153
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTPAsnTyrAsnThrAsnIleThrGluGluAsn 40
 DB 154 CTGTCTATCAAGTCACTTGGTCTTGGAAATTAAACCAATTTCTGAAGAGAT 213
 QY 41 ValGlnAsnMetAsnAlaGlyAspLysTTPSerAlaPheLeuLysGluGlnSerThr 60
 DB 214 GTCCAAAGATGATATCTGCGGACAAATGTCCTTTTAAAGGACAGTCCACA 273
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 274 CTGGCCAAATGATCATCTACAGAAATTCAGAAATCTCACAGTCAAGCTTCAAGCTGCG 333
 QY 81 AlaLeuGlnGlnAsnLysSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 334 GCTCTCAGCAAAATGGGCTTCAGTCTCAGAGACAGCAAGCAAGCGTTGAACACA 393
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnPro 120
 DB 394 ATTCATAATCAATGAGCACATCTACACTGGAAGAGTTGTAACCCAGATATATCA 453
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 454 CAAGAAATGCTTATCTTACCTGACCAAGGTTGAATGAATATATGCAACAGTTTACACTAC 513
 QY 141 AsnGluArgLeuTyrPalaTyrPalaTyrPalaTyrPalaTyrPalaTyrPalaTyrPala 160
 DB 514 AATGAGAGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGGTGGGCAAGAGCTGAGGCCA 573
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 574 TATATGAAGAGATGTGCTTGAATAATGAGATGCAAGAGCAAAATCATATATGAGGAC 633
 QY 181 TyrGluAspTyrTyrPalaTyrPalaTyrPalaTyrPalaTyrPalaTyrPalaTyrPala 200
 DB 634 TATGGGATTTTGGAGAGAGAGACTTGAAGATGAGGATGAGGCTATGATGATGAGC 693
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 DB 694 CGGGGCACTGATGATGAGATGAGAACTATGAAATGAGGATGAGGCTATGATGATGAGC 753
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 754 CATCTTCATGCTATGCTGAGGCAAGTGAATGAATGAGGCTATGCTTCATGCTGAGTCA 813
 QY 241 IleGlyCysLeuProAlaHisLeuGluLysAspMetTyrGlyArgPheThrAsnLeu 260

DB 814 ATGGATCCCTCCCTGCTCATTTGCTTGTGCTATATGCGGCTAGATTTTGGACAAATCTG 873
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 874 TACCTTTGAGAGTTCCTCTTGGAGAGAAACCAATGATGTTACTGATGCAATGCTG 933
 QY 281 AspGlnAlaTyrPalaGlnArgIlePheLysGluAlaGlyLysPhePheValSerVal 300
 DB 934 GACAGGCGCTGGAGTGCACAGAGATATTCAGAGAGCGCGAGAGTCTTGTATCTGTT 993
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrPalaLysSerMetLeuThrAspProGlyLys 320
 DB 994 GGTCTCTTATATGATCACTCAAGATTTCTGGAAATTCATCTATACGAGACCCAGGAAT 1053
 QY 321 ValGlnLysAlaValLysHisProThrAlaTyrPalaLysGlyLysGlyAspPheArgIle 340
 DB 1054 GTTCAGAAAGAGTCTGCTGATCCACAGCTTGGGAGCTGGGAGAGGCGAGCTTCAAGATC 1113
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHis 360
 DB 1114 CTTATGTCACAAAGTGCACATGACACACTTCTGACAGCTCATATGAGATGGGCAT 1173
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1174 ATCCAGTATGATATGCTATATGCTGACACACCTTTCTCTTAAAGAAATGGAGTAAATGA 1233
 QY 381 GlyPheHisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1234 GGATTCATGAGAGCTGTTGGGAAATCATGTCACCTTCTGAGGACACACCTTAACATTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 DB 1294 AAATCCATGCTGCTCTGCTCACCAGATTTTCAAGAGCAATGAAAGAAATTAACCTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1354 CTGCTCAACAGACACTACATGTTGGGCTGCTGCCATTTACTTACATGTTTGAAGAAG 1413
 QY 441 TrpArgTyrMetValPheLysGlyLysIleProLysAspGlnTyrMetLysLysTyrTrp 460
 DB 1414 TGAGAGTGAAGTGTCTTAAAGGGGAATTCACAAAGCAGCTGAGTGAAGAAATGGTGTG 1473
 QY 461 GluMetLysArgGluIleValGlyValValLysProValProHisAspGluThrTyrCys 480
 DB 1474 GAGATGAAGCAGAGATGATGTTGGGTGTCAGCTGCGCCCATGATGAACACTACTGT 1533
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrArg 500
 DB 1534 GACCCGCACTCTGCTCATGTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1593
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 DB 1594 ACCCTTACCAATTCACATTTCAAGAGACACTTTGTCAAGAGAGCTAAACATGAAGGCGCT 1653
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyLysLysLeuPhe----- 537
 DB 1654 CTGCACAAATGTGACATCTCAACCTACAGAAAGCTGAGAGAAAGCTT- GTAAGAAAT 1712
 QY 537 ----- 537
 DB 1713 ACCTCAAAATGTGAACCTCTCTAGATTTCAATTTCAATTTCAATTTCAATTTCAATTTCA 1772
 QY 537 ----- 537
 DB 1773 TATTGATTTCTTTCTCTAAAGAAATTTTATGGCTCAAAATGCTCATTTTACAA 1832
 QY 537 ----- 537
 DB 1833 ACCAAACATTAATTTGTGTGACAGAGAACTAGAACATCAACAATTTGGGTGGCCCA 1892
 QY 537 ----- 537
 DB 1893 CCTCTTTCTCCCTATCACTAATCAAGCCCTCTCTCTGATTAATGGAAGAAAGAGCG 1952

```

QY 537 ----- 537
Db 1953 GTTAAAGGTGGAATATCTGTTAAATGACATCTTTCTTATCTGCCAGAAACAATTT 2012
QY 537 ----- 537
Db 2013 AGCCAAAGTCAAGAGAGAAACCATGATCATGATGTAATATATGTACATCTGGAAACC 2072
QY 538 -----AsnMetLeuArgLeuGlyLysSer 545
Db 2073 CCTCAAAAGGCCCTGACCCCTTTTGTGTGACAAATATCTGAGCTTGGAAATATCA 2132
QY 546 GIUPTPTPTThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgPro 565
Db 2133 GAACCTGGACCTTGCATGATGGAATATGTTAGAGC-AGAATCATGAATGTAAGGCUCA 2191
QY 566 LeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGluAsnLysAsnSerPhe 585
Db 2192 CTGCTCACTACTTTGAGCCCTTATTACCTGGCGGAAGACCAAGACAAAGATTTCTTT 2251
QY 586 ValGlyTyrSerThrAspTyrSerProTyrAlaAsp-GlnSerIleLysValArgIleSe 605
Db 2252 GTGGGATGAGTACCGACTGGAGTCCATATGCAGACCAAAAGCATCAAAAGTGAAGTAA 2311
QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPh 625
Db 2312 CCTAAATCAGCTCTTGGAGATGAAGCATATGATGAGACAAATGAATGTAATCTCTTAA 2371
QY 625 eArgSerSerValAlaTyrAlaMetArgLntyrPheLeuLysValLysAsnGlnMetI 645
Db 2372 CCGATCATCTGTTGATATGCTATGAGCAGTACTTTTAAAGTAAATATCATGATGAT 2431
QY 645 eleuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
Db 2432 TCTTTTGGGGAGGAGATGTGCGAGTGAATTTGAACCAAGAAATCTCTTAAATTT 2491
QY 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgTrpArgValGluLysAl 685
Db 2492 CTTTCTCATCTGACCTTAAATATGCTCTGATATCTTCTTGAACCTGAAAGTGAAGAGGC 2551
QY 685 aIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuG 705
Db 2552 CATCAGATGTCGCCGAGCCGATCATGATGATGCTTCCGCTGATATGACCAACGCTTGA 2611
QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
Db 2612 GTTCTGGGAGTACAGCCACCTTGGACCTCTCAACGAGCCCT 2657

RESULT 8
US-10-121-049-71
; Sequence 71, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME

```

```

; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 71
; LENGTH: 3732
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-121-049-71

Alignment Scores:
Pred. No.: 0
Score: 3720.00
Percent Similarity: 83.88%
Best Local Similarity: 83.76%
Query Match: 96.15%
DB: 9
gaps: 1

US-09-978-385-2_COPY_19_738 (1-720) x US-10-121-049-71 (1-3732)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
Db 94 TCACACATGAGAGACAGGCCAAGACATTTTGGACAAGTTTACCAACGACGAGAC 153
QY 21 LeuPheTyrGlnSerSerIleuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsn 40
Db 154 CTGTTCTATCAAGTTCACCTCTGCTGGAATTAATACCAATATTAATCTGAAGAGAA 213
QY 41 ValGlnAsnMetAsnAlaGlyAspLysTrpSerLysPheLeuLysGluGlnSerThr 60
Db 214 GTCCAAACATGATATATGCTGGGACAAATGCTGCTTTTAAAGAACGTCACACA 273
QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
Db 274 CTTCGCCAAATGATATCCACTACAAAGAAATTCAGAACTCAGCTCAAGCTCAGCGAG 333
QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
Db 334 GCCTTCAGCAAAATGGCTCTTCACAGTCTCAGAGACAAAGCAAGCGTTGAACACA 393
QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAsnPro 120
Db 394 ATTCTAATATCAATGACACCATCTACAGTCTGAGAAAGTTGTATACCAATATTCGA 453
QY 121 GlnGlnCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
Db 454 CAAGAAATGCTTATTAATGACAGGTTGAATGAATATATGGCAACAGTTTAACTAC 513
QY 141 AsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
Db 514 AATGAGAGGCTCTGGGCTTGGGAAACCTGAGATCTGAGAGCTGCGCAAGCAGCTGAGGCCA 573
QY 161 LeuTyrGlnGluTyrValValIleuLysAsnGluMetAlaArgAlaAsnHisLysTrpGluAsp 180
Db 574 TTATATGAAAGATATGCTGCTTGAATAATGATGCAAGCAATATATATATAGAGGC 633
QY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyLysAspTyrSer 200
Db 634 TATGGGATATATGAGAGAGAGACTATGAAGTAAATGGGCTGATGCTATGATACATACAGC 693
QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
Db 694 CGCGGCCAGTGTATTAAGATGGAACATGACATCTTGAAGAATTAACCATTAATATGAA 753
QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAlaAlaTyrProSerTyrTrpLysSerPro 240
Db 754 CATCTTCATGCTATATGAGGGAAGGAAATGATGAATGCTATCTTCATATATGAGTCA 813
QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheThrPheAsnLeu 260
Db 814 ATTGATGCTCTCCCTGCTCATTTGCTTGGTATATGTGGGATGATTTTGGCAAAATCTG 873

```


CURRENT FILING DATE: 2002-04-16
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 71
LENGTH: 3732
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-904-71

Alignment Scores:

Pred. No.:	0	Length:	3732
Score:	3720.00	Matches:	717
Percent Similarity:	83.88%	Conservative:	1
Best Local Similarity:	83.76%	Mismatches:	2
Query Match:	96.15%	Indels:	138
DB:	9	Gaps:	1

US-09-978-385-2_COPY_19_738 (1-720) x US-10-123-904-71 (1-3732)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGlnAlaGluAsp 20
DB 94 TCCACCATGAGGAAGAGCCAAAGCATTTTGGACAAAGTTTACCAGCAAGCCGAAAGC 153
QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheAsnTyrAsnThrAsnIlethGlnGluAsn 40
DB 154 CTGTTCTATCAAAAGTTCACTGCTCTTGGAAATTATACACCATAATTTACTGAAGAGAT 213
QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerAlaPheLeuLysGlnGlnSerThr 60
DB 214 GTCCAAACATGAATATATCTGGGCAAAATGTCCTGCTTTTAAAGAAACAGTCCACA 273
QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
DB 274 GTTGGCCAAATATCTCACTACAAAGAAATTCAGAAATCTCCAGTCAAGCTTCAGTGCAG 333
QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerTyrAsnThr 100
DB 334 GCTCTTCAAGCAAAATGCTCTTCACTGCTCTCAAGACAAAGACAAAGCGTTGACACA 393
QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyValCysAsnProAspAsnPro 120
DB 394 ATTTCAATATACATGAGCACCACATCTACAGTACTGAGAAAGTTTGAACCCAGATATCA 453
QY 121 GlnGluCysLeuLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
DB 454 CAAAGATGCTTATTTCTTGAACCAAGTTGAATGAATGAAGCAAAAGTTTACATAC 513
QY 141 AsnGluAspLeuTyrPheLysLeuSerTyrPheSerGluValGlyLysGlnLeuAspPro 160
DB 514 AATGAGAGGCTCTGGGCTGGGAAAGCTGGAGATCTGAGGTGCGCAACAGCTGAGGCCA 573
QY 161 LeuTyrGlnGluTyrValValLeuLysAsnGluMetAlaAsnHisTyrGlnAsp 180
DB 574 TTAATATGAAGATATGCTGCTTGAATAATGAGAGCAAGCAAGCAATTAATGAGAGAC 633
QY 181 TyrGluAspTyrTyrArgGlyAspTyrGluValAsnGluValAspGlyTyrAspTyrSer 200
DB 634 TATGGGATATTTGAGAGAGAGACATATGAAGTAAATGGGGTAGAGGCTATGACTACAGC 693
QY 201 ArgGluGlnLeuIleGluAspValGluHisThrPheGlnGluIleLysProLeuTyrGlu 220
DB 694 CGCGGCCAGGTGATGAGAGATGAGACATACCTTTGAAGATTAACCATTTATATCA 753
QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
DB 754 CATTTCTCAATGCTATGATGAGGCAAAAGTTGAATGAAATCCATCTATATACAGTCCA 813
QY 241 IleGlyCysLeuProAlaHisLeuLeuLysAspMetTyrGlyArgPheTyrThrAsnLeu 260
DB 814 ATTTGATGCTCCCTGCTCATTTGCTTGTGATATGTTGGGTAGATTTTGAACAAATCTG 873
QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValIleAspAlaMetVal 280

DB 874 TACTCTTTGACAGTTCCTTTGGACAGAAACCAACATAGATTTACTGATGCAATGCTG 933
QY 281 AspGlnAlaTyrPheAlaGlnArgIlePheLysGluIleGlyLysPhePheValSerVal 300
DB 934 GACCAAGGCTTGGATGACACAGAAATATTCAGAGAGGCGCAAGATTTCTTGTATCTGTT 993
QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrPheLysAsnSerMetLeuThrAspProGlyAsn 320
DB 994 GGTCTTCTATATGATGACATGAGATTTCTGGGAAATTTCCATGCTAACAGCAACAGAAAT 1053
QY 321 ValGlnLysAlaValCysHisProThrAlaTyrPheLysGlnLysGlyAspPheArgIle 340
DB 1054 GTTCAGAAAGCAGTCTGCATCCACAGCTTGGGCAAGCTGGGGAAGGCGACTTCAGAGATC 1113
QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrLanHisGlnMetGlyHis 360
DB 1114 CTATATGTCACAAAGGAGGACATGACACTTCTGACAGCTCATCTAGATGAGGAGCAT 1173
QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
DB 1174 ATCCAGATGATATGATGATATGCTGACAAACCTTTCTGCTAAGAAATGAGCTAATGAA 1233
QY 381 GlyPheHisGlnAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
DB 1234 GGATTCATGAAGCTGTTGGGAAATCATGTCACTTCTCCACCCACCTTAAGCATTTTA 1293
QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
DB 1294 AAATCCATTTGCTCTGTCACCCGATTTTCAAGAAACATGAACAGAAATTAACCTTC 1353
QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
DB 1354 CTGCTCAAAACAGCAGCTCAGATGTTGGAGCTCTGCACTTACTTACATGATTAGAGAG 1413
QY 441 TyrArgThrMetValPheLysGlyGluIleProLysAspLysIlePheLysTyrPhe 460
DB 1414 TGGAGGTGAGGCTTTTAAAGGGGAAATTCACAAACCCAGTGGATGAAAGAGTGTGG 1473
QY 461 GluMetLysArgGluIleValGlyValGlyProValProHisAspGluThrTyrCys 480
DB 1474 GAGATGAAGGAGAGATAGTTGGGGTGGTGGAACTGTGGCCCATGATGAACATATCTGT 1533
QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
DB 1534 GACCCGCACTCTGTTCCATGTTTCTGATGATTAATCAATTCATGATATTCACAAAG 1593
QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaIleHisGlnGluPro 520
DB 1594 ACCCTTACCAATTCAGTTTCAAGAAACCTTTGTCAAGCAGCTTAACATGAAGGCCCT 1653
QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe 537
DB 1654 CTGCACAAATGTGACATCTCAACCTCTACAGAAAGCTGAGACAGAAACTGTT-GRNAAGAAAT 1712
QY 537 537
DB 1713 ACCTCAAAATGTTGAACCTCTCTAGTATTCAGTATTAATTCATTCATTCAGCTAGGTTTG 1772
QY 537 537
DB 1773 TATTTGATTTCTTTGTTCTAATAAAGAAATTTTATGCGCTCAAAATGCTCTCATTTACA 1832
QY 537 537
DB 1833 ACCAAACATTTAATTTGTGTGTCAGACAGAACTAGACATACAAACATTTGGGTGGCCA 1892
QY 537 537
DB 1893 CCTCTTTTCTCCATCATTAATACAGCCCTCTCTCTGTAATTTGAAGAAAGAGCG 1952
QY 537 537
DB 1953 GTTTAGGCTGGAATATATCTGTTAATATGCAATCTTTTCTTATCTGCCAGAAAGCAATTT 2012

QY	537	-----	537
Db	2013	AGCCAAAGTCAAAAGGAAAGAAACCATAGATCTATAGTGTAAATATATATGTACATCTGGAAAC	2072
QY	538	-----	-----
Db	2073	CCTCAAAAGGCCCTGAAGACCCCTTTTWTWTGTAAGCAATATGTGAGGCTTGAAAAATCA	2132
QY	546	GIUROTPTPTHTLLEuLlALEuGLlASuValValGLlAlAluysaAnMeLAsuValArgPro	565
Db	2133	GAACCCGTGGACCCCTGAGATTTGGAAAGTGTGTAGAGAC - AAGAACATGAAAGTGAAGCCA	2191
QY	566	LeuLeuSuSuTYrPheGLuProLeuPheThrTYrPheuLysAspGlnAsuLysAsnSerPhe	585
Db	2192	CTGCTCAACTACTTTTGAGCCCTTTATTTACCTGGCTGAAAGACCCGAAACAATATCTTTT	2251
QY	586	ValGLlTYrPserThrAspTYrPserProTYrAlaAsp - GlnSerlLeuValArgLese	605
Db	2252	GTGGAGATGAGATACCGACTGGAGTGCATATGACAGACCCAAAGCATCAAGTGAAGATPAG	2311
QY	605	rLeuLysSerLAlALEuGLlAspLysAlATyGLlITrPAsnAspSngLlMeTYrLeuPh	625
Db	2312	CCCTAAACACACCTCTTGAGATPAAAGCATTAAGATGAGACGACATGAAAGTACCTGTT	2371
QY	625	eArgSerSerValAlATyAlaMeTArgGLnTYrPheuLysValLysAsnGLlMetLl	645
Db	2372	CCGATCACTCTTGATCATCTGTATGAGGACATCTTTTAAAGTMAAAATTCAGATAT	2433
QY	645	eLeuPheGLlYGLlGLlAspValArgValAlAlAsnLeuLysProArgLesePheAsnPh	665
Db	2432	TCTTTWTGGGAGAGAGATGTGCGAGTGGCTAAATTGTGAACCAAGAAATCTCTTAAATTT	2491
QY	665	ePheValThrAlaProLysAsnValSerAspLlLeLeProArgThrgLlUValGLlULysAl	685
Db	2492	CTTTGTGCCTCACCCTAAATAAATGTCTGATATCATCTTCCTGAAGTGAAGTGAAGAACGC	2551
QY	685	AlLeuArgMetSerArgSerArgLlLeAsnAspAlAPheArgLeuAsnAspAsnSerLeuGLl	705
Db	2552	CATCAGATGTGCCGGAGACCGTATCAATAGATGCTTCCGTGTGAATGACAACAGCTGAGA	2611
QY	705	uPheLeuGLlYlLeGLlProThrLeuGLlProProAsnGLlProPro	720
Db	2612	GTTCCTGGGATACAGCCACACTTGGACCTCTCAACAGCCCTT	2657
RESULT 10			
US-10-140-470-71			
; Sequence 71, Application US/10140470			
; Publication No. US2003002231A1			
GENERAL INFORMATION:			
; APPLICANT: Baker, Kevin P.			
; APPLICANT: Beresini, Maureen			
; APPLICANT: Desrogers, Laura			
; APPLICANT: Desnoyers, Luc			
; APPLICANT: Filvaroff, Ellen			
; APPLICANT: Gao, Wei-Qiang			
; APPLICANT: Gerritsen, Mary E.			
; APPLICANT: Goddard, Audrey			
; APPLICANT: Godowski, Paul J.			
; APPLICANT: Gurney, Austin L.			
; APPLICANT: Sherwood, Steven			
; APPLICANT: Smith, Victoria			
; APPLICANT: Stewart, Timothy A.			
; APPLICANT: Tumas, Daniel			
; APPLICANT: Watanabe, Colin K.			
; APPLICANT: Wood, William			
; APPLICANT: Zhang, Zemin			
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
FILE REFERENCE: P3330R1C160			
CURRENT APPLICATION NUMBER: US/10/140,470			
CURRENT FILING DATE: 2002-05-06			
Prior Application removed - See Palm or File Wrapper			

```

; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 71
; LENGTH: 3732
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-470-71

Alignment Scores:
Pred. No.: 0
Score: 3720.00
Percent Similarity: 83.88%
Best Local Similarity: 83.76%
Query Match: 96.15%
DB: 9

Length: 3732
Matches: 717
Conservative: 1
Mismatches: 2
Indels: 138
Gaps: 1
US-09-978-385-2_COPY_19_738 (1-720) x US-10-140-470-71 (1-3732)

```

OY	SetThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp	20
Db	94 TCACCACTTAGAGAAACAGGCCAAGACATTTTGGCAACCTTTAACCCAGAACCCGAAGAC	15
OY	21 LeuPheIleYrGlnSerSerLeuAlaSerTrpAsnTrpAsnIleThrGluGluAsn	40
Db	154 CTGATCTATCAAGATTCACCTGCTTCTTGATATTTAAACCAATATTACTGAAGACAAT	213
OY	41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluInSerThr	60
Db	214 GTCCAAACATGAAATATATGCTGGGGCAAAATGCTGCTTTTAAAGCAACAGCTCCACA	273
OY	61 LeuAlaGlnMetYrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGluInGln	80
Db	274 CTTCGCCAAATGATATCCACTACAGAAATTCAGAAATTCACAGTCACAGCTTCACCTCGAC	333
OY	81 AlaLeuGlnGlnAsnGlySerValLeuSerGluAspLysSerLysArgLeuAsnThr	100
Db	334 GCCTTCACGCAAAATGGTGCTTCAGCGCTCAGACAGCAACAGCAACAGCGTTGACACA	393
OY	101 IleLeuAsnThrMetSerThrIleYrSerThrGlyLysValCysAsnProAspAsnPro	120
Db	394 ATTCTAAATATCAATGAGCACACCATCTACAGTACTGGAAAAAGTTGTAAACCCAGATATCCA	455
OY	121 GlnGluCysLeuLeuLeuGluProGluLysAsnGluIleMetAlaAsnSerLeuAspTyr	140
Db	454 CAAGAATGCTTATTACTGTACCAACGAGTTTGAAATGAATATGCAACAGTTTGAAGCTAC	513
OY	141 AsnGluArgLeuThrAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro	160
Db	514 AATAGAGGCTCTGGGGCTTGGAACCTGGAGATCTGAGGTGGCAACGACGTGAGGCCA	573
OY	161 LeuTyrGluGluIuYrValIleLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp	180
Db	574 TTATATGAAGATGATGTGGTCTTGAAAATGAGATGGCAAGGCAATCATTTAGAGGAC	633
OY	181 TyrGlyAspTyrTrpArgLysAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer	200
Db	634 TATGGGATATTATGGAGAGAGACTATGAAGTAAAGGGGTAGATGGCTATGACTACAGC	693
OY	201 ArgGlyLeuIleuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu	220
Db	694 CGGGCCAGTTGATGTAAGATGTGGAAACATCCTTTGAAGACAATTAACCAATTAATATAA	753
OY	221 HisLeuHisAlaTrpValArgAlaLysLeuMetCAsnAlaTyrProSerTyrIleSerPro	240
Db	754 CATTTTCATGCTATGTGAGGGCAAGATTGATGAATGCTATTCCTTCCATATCAGTCCA	813
OY	241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeu	260
Db	814 ATTTGAATGCTCCCTCGCTCATTTGCTTGAGATATGTGGGAGATGATTTTGGCAAAATCTG	873
OY	261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal	280
Db	874 TACTCTTTGACAGTTCCTCTTGACAGAAACCAACATATGATGTTTACATGATCAATGGTG	933

QY 281 ASFGINAlATPASPAlaGlnArgrIlePheIysGlnAlaGlnIysPhePheValSerVal 300
 DB 934 GACCAGGCGTGGGATGCCACAGAGATATTCAGAGAGCGGAGAGGTTCTTTGTATCTGCTT 993
 QY 301 GlysLeuProAsnMetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGlyAsn 320
 DB 994 GGTCTTCCTCAATATGACTCAAGAGATTCGGGAAAAATTCATCGTACAGGACCCAGGAAT 1053
 QY 321 ValGlnIysAlaValAlaCysHisProThrAlaTrpAsnLeuGlyIysGlyAspPheArgIle 340
 DB 1054 GTTCAGAAAGAGCTGGCCATCCACACCTGGGAGACCTGGGAGGCGGACTTCAGATC 1113
 QY 341 LeuMetCysThrIysValThrMetAspAspPheLeuThrAlaHisIleGlnMetGlyHis 360
 DB 1114 CTATAGTCACAAAGAGTACATGACAGCTCTTCGAGCTCATTCATGAGATGGGCAT 1173
 QY 361 IleGlnIysAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1174 ATCCAGTATGATATGGCATATGTCGACCAACCTTTCTGTAAGAAATGGAGCTAATGAA 1233
 QY 381 GlyPheHisGlnAlaValAlaGlyGlnIleMetSerLeuSerAlaAlaThrProIysHisLeu 400
 DB 1234 GGATTCATGAGAGCTGGTGGGAAATCATGTCACCTTCGCGACCACTTAAGCATTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGlnAsnGlnIysProIysHisLeu 420
 DB 1294 AAATCCANTGCTCTCTGTCACCGATTTTCAGAAACATGAAACAGAAATTAACCTTC 1353
 QY 421 LeuLeuIysGlnAlaLeuThrIleValAlaGlyThrLeuProPheThrIleLeuGlnIys 440
 DB 1354 CTGCTCAACAAACAGCAGCATGTTGGGACTCTGCATTTCTGATCATGTTAGAGAAC 1413
 QY 441 TrpArgTrpMetValPheIysGlyGlnIleProIysAsnGlnIlePheIleIysTrpTrp 460
 DB 1414 TGGAGGGAGAGCTGCTTAAAGGGGAAATCCCAAAACAGCGAGGAGGAAAGTGCTGG 1473
 QY 461 GluMetLysArgGlnIleValAlaGlyValAlaGlnProValProHisAspGlnIleAsnPhe 480
 DB 1474 GAGATGAGAGGAGAGATAGTTGGGCTGTGGAACTGTGCCCATGATGAAACATCTGCT 1533
 QY 481 AspProLaserLeuPheHisValSerAsnAspIleSerPheIleArgIleIleArgIle 500
 DB 1534 GACCCCGCATCTGCTCCATGTTTCGATGATGATGATGATGATGATGATGATGATGAT 1593
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaIysHisGlnIlePro 520
 DB 1594 ACCCTTACCAATTCAGTTTCAAGAGCAGCTTGTCAACGACCTTAACGATGAAGCCCT 1653
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyGlnIysLeuPhe 537
 DB 1654 CTGCACAAATGTGACATCTCAAACTCTACGAAAGCTGGACAGAACTGTT-GTAGAAT 1712
 QY 537 537
 DB 1713 ACCTCAAAATGTTGAACCTCTCCTAGTATTCATGATTCATGCTAGGTG 1772
 QY 537 537
 DB 1773 TATTGATTTCTTTGTTCTAAAGAAATTTTATGCGCTCAAAATGCTCATTTACA 1832
 QY 537 537
 DB 1833 ACCAAACATTATTTGTGTCAGACAGAGACCTACATACATACATATGGGTGGCCA 1892
 QY 537 537
 DB 1893 CCTCTTTTCCCTATCATACACAGCCCTCTCTCTGTAATGTAAGAAAGAGAGCG 1952
 QY 537 537
 DB 1953 GTTTAGGAGTGAATATATCTGTTAATATGCAATCTTTTCTGTCGAGAGCAAAATTT 2012
 QY 537 537

DB 2013 AGCCAACTCAAGAGAGAAACCATGATCATGATGTAATATATGATCATCTGGAACC 2072
 QY 538 -----AsnMetLeuArgLeuGlyIysSer 545
 DB 2073 CCTCAAAAGCCCTGAACCCCTTTTGTGTAGCAAAATATGCTGAGCGCTGGAAATCA 2132
 QY 546 GluProTrpThrLeuAlaLeuGlnAsnValAlaGlyAlaIysAsnMetSerValArgPro 565
 DB 2133 GAACCGTGGACCTTACATGTTGAAATGTTGTAGAGCC-AGAAACATGATGTAAGGCCA 2191
 QY 566 LeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuIysAspGlnAsnLysAsnSerPhe 585
 DB 2192 CTGCTCAACTACTTTAGCCCTTATTTACCTGGCTGAAGACCAAGCAAGAAATCTTTT 2251
 QY 586 ValGlyTyrSerThrAspTrpSerProIysAlaAsp-GlnSerIleIysValArgIleSe 605
 DB 2252 GTGGGATGGAGTACGAGCTGAGTCCATATGACAGACCAACCAACATCAAGATGAGTAAG 2311
 QY 605 IleLysSerAlaLeuGlyAspLysAlaTyrGlnIysProAsnAspAsnGlnMetTyrLeuPh 625
 DB 2312 CCTAAATACAGCTCTTGAGATTAACATATGATGAAACGCAATGAAATGTACCTGT 2371
 QY 625 eArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuIysValIysAsnGlnMetIle 645
 DB 2372 CCGATCATCTGTTGATATGATGATGAGCAGTACTTTTAAAGTAAATCAATCATGAT 2431
 QY 645 eLeuPheGlyGlnGlnAspValArgValAlaAsnLeuIysProArgIleSerPheAsnPh 665
 DB 2432 TCTTTTGGGAGAGAGATGTGCGAGTGCATTAATTTGAACCAAGAACTCTCTTAATTT 2491
 QY 665 ePheValThrAlaProIysAsnValSerAspIleIleProArgTrpGlnValAlaGlnIysAl 685
 DB 2492 CTTTGTCACTGCACCTAAATAATGTGTGATATATCTCTGAACCTGAAGTTGAAAGGC 2551
 QY 685 aIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAsnSerLeuGln 705
 DB 2552 CATCAGAGAGTCCCGAGCGGATCAATGATGCTTCCTGCAATGACACAGAGCTAGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2612 GTTTCGGGAGATACAGCCACACTTGACCTCTCAACAGCCCTT 2657
 QY 720 720
 DB 2657 2657
 RESULT 11
 US-10-175-746-71
 ; Sequence 71, Application US/10175746
 ; Publication No. US20030027270A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C353
 ; CURRENT APPLICATION NUMBER: US/10/175,746
 ; CURRENT FILING DATE: 2002-06-19
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 71


```

; LENGTH: 3732
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-746-71

Alignment Scores:
Pred. No.: 0          Length: 3732
Score: 3720.00       Matches: 717
Percent Similarity: 83.88%    Conservative: 1
Best Local Similarity: 83.76%  Mismatches: 2
Query Match: 96.15%          Indels: 138
DB: 9                  Gaps: 1

US-09-978-385-2_COPY_19_738 (1-720) x US-10-175-746-71 (1-3732)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
Db 94 TCACACATTGAGGAGACAGGCAAGACATTTTGGACAACTTTAACACAGCAACCGAGAGC 153
QY 21 LeuPheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsn 40
Db 154 CTGTCTTATCAAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 213
QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
Db 214 GTCCAAACATGATATATGCTGGGACAAATGCTGCTTTTAAAGGACAGTCCACA 273
QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
Db 274 CTGGCCAAATGATCCACTACAAAGAAATTCAGAAATCTCACAGTCAAGCTCAGCTCAG 333
QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
Db 334 GCTCTTCACGAAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 393
QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
Db 394 ATCTCAATATACATGAGCAACATCTACACTGAGAAAGTTTGTAAACCACTAATACCA 453
QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
Db 454 CAAGATGCTTATTAATCTGACCAAGCTTGAATGAATGAATGAATGAATGAATGAATGA 513
QY 141 AsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuAlaPro 160
Db 514 AAGGAGAGGCTCTGGGCTGGGAAAGCTGAGATCTGAGCTGGGCAAGCAAGCTGAGCCA 573
QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
Db 574 TTATATGAGAGATATGCTGCTTGAATAATGAGATGAGATGAGATGAGATGAGATGAG 633
QY 181 TyrGlyAspTyrTrpArgLysAspTyrGluValAsnGlyValAspLysTyrAspTyrSer 200
Db 634 TATGGGGAATATGAGAGAGAGACTATGAATGAATGAGGATGAGATGAGATGAGATGAG 693
QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
Db 694 CGGGGCGAGTGTATTAAGATGTGAACATACCTTGAAGAGATTTAAACCATTTATATGAA 753
QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
Db 754 CATCTTCATGCTATGAGAGGCAAAAGTTGATGAATGCTTCTCTCTCTCTCTCTCTCTCT 813
QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpPylarPheTrpThrAsnLeu 260
Db 814 ATTGAGAGCTCTCCCTGCTATTTGCTGGATATGAGGATGAGATTTTGAACAAATCTG 873
QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
Db 874 TACTCTTTCAGACTCTCTCTTGGACAAACCAACATAGATGCTCTAGTGAAGAGGCTG 933
QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300

```

```

Db 934 GACCAGGCTGTGGATGACAGAGATATATCAAGGAGGCCGAGAGTCTTTGTATCTGTT 993
QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
Db 994 GGTCTTCTTATATGACTCAAGGATTCGGGAAATTTCCATGATCAAGGACCCGAGAAAT 1053
QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
Db 1054 GTTCAGAAAGAGCTGTGCATGCCACAGCCTTGGGAGCTGGGAGGCGAGCTTCAGATC 1113
QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisLysGluMetGlyHis 360
Db 1114 CTATATGTCACAAAGTGCATCAATGAGACACTCTCTGACAGCTCATCATGATGGGCAT 1173
QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
Db 1174 ATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1233
QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
Db 1234 GATTCATGAGAGCTGTGGGAAATCATGTCACCTTCTGACGCCACACCTTAACCATTTA 1293
QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
Db 1294 AATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1353
QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
Db 1354 CTGCTCAAAACAGCACTCAGATGTTGTCAGCTGTCATTTACTTACATTTAGAGAAAG 1413
QY 441 TrpArgTrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTrp 460
Db 1414 TGAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1473
QY 461 GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCys 480
Db 1474 GAGATGAGGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1533
QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
Db 1534 GACCCCGCATCTCTGTCATGTTCTGATGATATCTCATTCATTCATTCATTCATTCAT 1593
QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGluPro 520
Db 1594 ACCCTTTACCAATTCAGATTCAGAAACACTTGTCTCAAGAGCTTAAACAGAAAGGCT 1653
QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyLysLysLeuPhe----- 537
Db 1654 CTGCACAAATGTGACATCTCAACTCTACAGAGCTGACAGAAAGTGT-CTAAGAAAT 1712
QY 537 ----- 537
Db 1713 ACCCTAAATGTGAACCTCTCTAGTATTCAGTATTAATTAATTCATGCTAGGTTTG 1772
QY 537 ----- 537
Db 1773 TATTTGATTTCTTTGTCTTAAAAAGAAATTTATGCGCTCAAAATGCTCATTTACAA 1832
QY 537 ----- 537
Db 1833 ACCCAACATTTATTTGTGTCAGACAGGAACCTGACATACAAACATTTGGGTGGGCA 1892
QY 537 ----- 537
Db 1893 CCTCTTTCTCCCTATCATTAATACAGCCCTCTCTCTGTAATTTGAAAGAAAGAGCG 1952
QY 537 ----- 537
Db 1953 GTTTAGGTTGAGATATATCTGTTAATATGATTCCTTTCTTATCTCCAGAAAGCAATTT 2012
QY 537 ----- 537
Db 2013 ACCCAAGTCAAGAGAAAGAAACCATAGATCATAGATGTAAATATATGTACATCTGAAAC 2072

```


QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGluAsn 320
 |||||
 Db 994 GGGCTTCCTAAATATGACTCAAGATTCGGGAAAATTCATGCTAAGCAGACCAAGAAAT 1053
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 |||||
 Db 1054 GTTCAGAAAGACATCTGGCCATCCACAGCTTGGGACGTGGGAGAGGCGACATTCAGAGATC 1113
 QY 341 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisHisGlnMetGlyHis 360
 |||||
 Db 1114 CTTATGCGCAAGGAGCAATGAGACAGCTTCACAGCTCATCATGAGATGGGGCAT 1173
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 |||||
 Db 1174 ATCCAGTATGATATGGCAATATGCTGCACAACTTTCTGCTAAGAAATGAGCTAATGAA 1233
 QY 381 GlyPheHisGluAlaValAlaGlyLuiLemSerLeuSerAlaAlaThrProLysHisLeu 400
 |||||
 Db 1234 GGAATTCATGACCTTGTGGGAAATCATGCTCTTTCGACGCCACACTTAAGCATTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPh 420
 |||||
 Db 1294 AAATCCATGCTGCTCTGTCACCCGATTTTCAAGAGACATGAAACAGAAATAACTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlyLys 440
 |||||
 Db 1354 CTGCTCAACACACACCTCAGATGTTGGGACCTCCATTTACTTACATGTTAGAGAG 1413
 QY 441 TrpArgTrpMetValPheLysGlyLuiLemProLysAspGlnTrpMetLysLysTrp 460
 |||||
 Db 1414 TGGAGGTGGATGCTCTTAAAGGGGAAATTCACCAAGACACAGTGGATGAAAAAGTGTG 1473
 QY 461 GluMetLysArgLuiLemValGlyValValGluProValProHisAspGluThrTyrCys 480
 |||||
 Db 1474 GAATATGAAGGAGAGATGAGTGGGGGTGGGAACCTGCCCCATGATGAACATPACTGT 1533
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheLeuArgTyrTrpArg 500
 |||||
 Db 1534 GACCCCGCATCTGTGCTCAATGTTCTGATTTCTCATTTCTGATTTGATTAACACAAAG 1593
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 |||||
 Db 1594 ACCCTTACCAATTCAGATTTCAGAAAGACCTTTCAGACGCTTAACATGAGGCCCT 1653
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyLuiLysLeuPhe----- 537
 |||||
 Db 1654 CTGCACAAATGTGCATCTCAAACTCTACAGAAAGTGCAGAAACTGTT-GTAAAGAAAT 1712
 QY 537 ----- 537
 Db 1713 ACCTCAAAATGTGAACCTCTCCTAGATTCAGTATTTACTCATTTCCAGCTAGGTTTG 1772
 QY 537 ----- 537
 Db 1773 TATTGATTTCTTGTGTTAAAAAGAAATTTTATGSGCTCAAAATGTCCTCATTTACA 1832
 QY 537 ----- 537
 Db 1833 ACCAAACATTTAATTTGTGTCAGACAGAACCTTACCATCAACAATTTGGTGGGCCA 1892
 QY 537 ----- 537
 Db 1893 CCTCTTTTCCCTTATCATTAACACGCCCTCTCTTCTGTTAATTTGAAGAAAGAGCG 1952
 QY 537 ----- 537
 Db 1953 GTTAGGGTGAATATATCTGTTAATATGATCTTTTCTTATCTGCCAGAAAGCAATTT 2012
 QY 537 ----- 537
 Db 2013 AGCCAAGTCAAGAGAGAAACCATAGATCATAGATGTAATATATGTACATCTGGAACC 2072
 QY 538 ----- 538
 AsnMetLeuArgLeuGlyLysSer 545

Db 2073 CCTCAAAAGGCCCTGAACCCCTTTTGTGTAGCAATATGAGGCTTGAAGAAATCA 2132
 QY 546 GluProTrpThrLeuAlaLeuGlnAsnValAlaGlyAlaLysAsnMetAspValArgPro 565
 |||||
 Db 2133 GAACCTGTGACCTTGCATTTGGAAATATTTGTGAGAGC-AGAACAATGAATGAAGGCCA 2191
 QY 566 LeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 585
 |||||
 Db 2192 CTGCTCAACTACTTTGACCCCTTATTTACCTGGCGTGAAGACCAAGCAACAATTTCTTTT 2251
 QY 586 ValGlyTrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSe 605
 |||||
 Db 2252 GGGGATGAGTACCGACACTGTGCTCATATGACAGACCCCAAGACATCAAAAGTGAAGATAG 2311
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPh 625
 |||||
 Db 2312 CCTAAATACAGCTCTTGGAGATTAAGCATATGATGGAACACATGAATATGACCTGTT 2371
 QY 625 eArgSerSerValAlaTyrAlaMetLargGlnTyrPheLeuLysValLysAsnGlnMetL 645
 |||||
 Db 2372 CCGATCATCTGTTGCATATGCTATGAGCAGACTTTTAAAGTAAATATCAGATGAT 2431
 QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
 |||||
 Db 2432 TCTTTTGGGGAGGAGATGTCGAGTGGCTAATTGAAACCAAGAAATCTCTTAATTT 2491
 QY 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGlyLysAl 685
 |||||
 Db 2492 CTTTCTACCTGCACCTTAAATATGTTCTGATATATCTTCTGAACTGAAGTTGAAAGGC 2551
 QY 685 alleaArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuG 705
 |||||
 Db 2552 CATCAGAGATGCCCGAGCCGATCATATGATGCTTCCGCTGTGATGACACACACCTAGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 |||||
 Db 2612 GTTCTGGGAGTACAGCCACACACTTGACCTCAACAGGCCCT 2657

RESULT 13
 US-10-176-921-71
 ; Sequence 71, Application US/10176921
 ; Publication NO. US20030027276A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroft, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C288
 ; CURRENT APPLICATION NUMBER: US/10/176,921
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 71
 ; LENGTH: 3732
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; US-10-176-921-71

Alignment Scores:

Pred. No.: 0 Length: 3732
 Score: 3720.00 Matches: 717
 Percent Similarity: 83.88% Conservative: 1
 Best Local Similarity: 83.76% Mismatches: 2
 Query Match: 96.15% Indels: 138
 DB: 9 Gaps: 1

US-09-978-385-2_COPY_19_738 (1-720) x US-10-176-921-71 (1-3732)

QY 1 SerThrIleGluGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 94 TCCACCATGGAGAACAGGCGCAAGACATTTTGGACAAAGTTTAACCCAGGAAGCCGAAAC 153
 QY 21 LeuPheTYGlnSerSerLeuAlaSerTrpAsnTrpAsnTrpAsnTrpAsnTrpAsn 40
 DB 154 CTGTTCTATCAAAAGTCACTTCTCTTGGATTTTAAACCAATATTACGAAAGAT 213
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
 DB 214 GTCCAAAACATGAAATGCTGGGCAAAATGCTGCTTTTAAAGAAACAGCCACA 273
 QY 61 LeuAlaGlnMetTrpProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 274 CTGGCCAAATGATATCACTACAGAAATTCAGAAATCTCACAGTCAAGCTTCACTGCGAG 333
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 334 GCTGTCAGCAAAATGGGCTCAGTCTGCTCAGAAAGCAAGAGCAAGCGTTGACACCA 393
 QY 101 IleLeuAsnTrpMetSerThrIleTrpSerThrGlyLysValCysAsnProAspAsnPro 120
 DB 394 ATTTCAATATACATAGACACCATCTACAGTACTGAAAGTTTGTAAACCCAGATTAATCCA 453
 QY 121 GlnGluCysLeuLeuLeuGluProGluLeuAsnGluIleMetAlaAsnSerLeuAspTrp 140
 DB 454 CAAAGATGCTTTTACTTGAACCCAGGTTGAATGAATGAATGGCAACAGTTTACACTAC 513
 QY 141 AsnGluAspGluTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuAspPro 160
 DB 514 AATGAGAGGCTCTGCGGCGGAAAGCTGAGATCTGAGGCGCAACACTGAGGCCA 573
 QY 161 LeuTrpGluGluTrpValIleuLysAsnGlnMetAlaArgAlaAsnHisTrpGluAsp 180
 DB 574 TTAATGAGAGATAGTGTGCTTGAATAATAGATGAGCAAGCAAAATCATTAAGAGAC 633
 QY 181 TyrGlyAspTrpTrpArgGlyAspTrpGluValAsnGlyValAlaAspGlyTrpAspTrpSer 200
 DB 634 TATGGGAGATTAATGGAGAGAGACATAGAAATGGGTAAGGCTATAGACTACAGC 693
 QY 201 ArgGlyGlnLeuIleGluAspValGlnHisThrPheGluGluIleLysProLeuTrpGlu 220
 DB 694 CCGCGCCAGTTCATGAGAGATGTCGAAATACCTTTGAAAGATTAACCATTAATAGAA 753
 QY 221 HisLeuHisAlaTrpValArgAlaLysLeuMetAsnAlaTrpProSerTrpIleSerPro 240
 DB 754 CATCTTCATAGCTATAGTAGGGCAAAAGTTGATGAATGCTTCTCTATATAGTCCA 813
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlnLysAspMetTrpGlyArgPheTrpThrAsnLeu 260
 DB 814 AATGGATGCTCTCCGCTCATTTGCTGTGATATAGTGGGATATTGGAACAAATCTG 873
 QY 261 TyrSerLeuThrValProPheGlyLysLysProAsnIleAspValThrAspAlaMetVal 280
 DB 874 TACTCTTTGACAGTTCCTTTGGACAAACAAACATAGATGTTACTAGATGCAATGCTG 933
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 934 GACCGAGGCTGGAGATGACAGAGATATTCAAGGAGCGGAGAAAGTTTGTGTATCTGTT 993
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320

DB 994 GGTCTCTCTATATAGTCAAGAGATCTCGGAAATTTCCATGCTAAACGAGCCAGGAAT 1053
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 DB 1054 GTTCAGAAACAGAGCTCTCCATCCACAGCTTGGGAGCTGGGAAAGGCGCACTTCAAGATC 1113
 QY 341 LeuMetCysThrLysValIlePheAspAspPheLeuThrAlaHisIleGluMetGlyHis 360
 DB 1114 CTATTGTGCAAAAGGAGCAAAATGAGACAGACTTCTGACAGCTCATCATGAGATGGGCAAT 1173
 QY 361 IleGluTrpAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1174 ATCCAGATGATATGAGCATATGCTGCAACAACTTTCTGCTAAGAAATGAGCTAATGAA 1233
 QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1234 GAGTTCATAGAGCTGTGGGAAATCATGCTACTTCTGACCCACCACTTAAGCATTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPro 420
 DB 1294 AATCCATTTGCTCTGTCACCCGATTTTCAAGAAAGCAATGAAACAGAAATTAATCTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTrpMetLeuGlyLys 440
 DB 1354 CTGCTCAAAACAGCACTCACAGATGTTGGAGCTCTGCTCATTTAATCTTAAGTAAAGAG 1413
 QY 441 TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysTrpTrp 460
 DB 1414 TGGAGTGGATGGCTTTAAAGGAAATTCGCAAAAGACAGTGAATGAAAAAGTGTGG 1473
 QY 461 GluMetLysArgGluIleValGlyValValGluProValProHisAspGluThrTrpCys 480
 DB 1474 GAGATGAGGAGAGATAGTGGGGTGGGAACCTGCGCCATGATAAACATCTCTGT 1533
 QY 481 AspProAlaSerLeuPheHisValSerAspTrpSerPheIleArgTrpTrpArg 500
 DB 1534 GACCCGCACTCTGTTCCATGTTCTGATGATTAATCACTATCTTGCATATTACACAG 1593
 QY 1594 ACCCTTACCAATTCCTGTTCAAGAAAGCACTTGTCAAGACACTTAACATGAAAGCCCT 1653
 QY 501 ThrLeuTrpGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 DB 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 537
 QY 1654 CTGACAAATGATGACATCTCAAACTGACGAAGCTGGACGAAGAACTGTT-GTAAAGAAAT 1712
 QY 537 ----- 537
 QY 1713 ACCTCAAATGTTGAACCTCTCTAGATTCAGTATTACATTTCCATGCCATAGCTTGG 1772
 QY 537 ----- 537
 QY 1773 TATTGATTTCTTGTCTTAAAAAAAATTTTATGCGCTCAAAATGCTCATTTACAA 1832
 QY 537 ----- 537
 QY 1833 ACCAAACATTAAATTTGTGTGTGTCAGACAGAACTAGACCATACAAATTTGGTGGCCA 1892
 QY 537 ----- 537
 QY 1893 CCGCTTTTCCGCTCATCATTAACACAGCCCTCTCTCTGTAATTGGAAGGAAGAGCG 1952
 QY 537 ----- 537
 QY 1953 GTTTAGGTTGAATATATCTGTAATATGATGATCTTTCTTATCTGACAGAACAATTT 2012
 QY 537 ----- 537
 QY 2013 AGCCAACTCAAGAGAGAAAGCAATAGATCATAGATGAATATATGATCATCTGGAACC 2072
 QY 538 ----- AsnMetLeuArgLeuGlyLysSer 545
 DB 2073 CTTCAAAGGCGCTGAACCCCTTTTGTGTGATGCAATATGCTGAGGCTTGGAAAAATCA 2132

QY 546 GluProtrpThrLeuAlaLeuGluAsnValValGlyAlaIysAsnMetAsnValArgPro 565
 |||||
 Db 2133 GAACCCGTGACCCAGCATGATGAAATGTTGTAGAAC -AAGAAATGATGATGAGCCCA 2191
 QY 566 LeuLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 585
 |||||
 Db 2192 CTGCTCACTACTTGTGAGCCCTTATTTTACTGCTGTAAGACCAAGAAAGAAATTTCTTT 2251
 QY 586 ValGlyTrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSe 605
 |||||
 Db 2252 GTGGATGAGATGCCAGTCCAGTCCATATGACAGACCAAGCATCAAGTGAAGATGAG 2311
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetLysLeuPhe 625
 |||||
 Db 2312 CCTAAATCAGCTCTTGGAGATTAAGCATATGATGAGACACATGAAATGTACCTGTT 2371
 QY 625 eArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetL 645
 |||||
 Db 2372 CCGATCATCTGTCATATGCTATGAGCGAGTACTTTTAAAGTAAAGTAAATGATGAT 2431
 QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
 |||||
 Db 2432 TCTTTTGGGAGAGAGATGCGAGTGGCTAATTGAAACCAAGATCTCTTATTT 2491
 QY 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAl 685
 |||||
 Db 2492 CTCTGTCACATGACCTAAATATGTCTGATATCATCTTAAAGTGAAGTTGAAAGGC 2551
 QY 685 aLeuArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLysL 705
 |||||
 Db 2552 CATAGAGATGTCGCGAGACCGTATCAATGATGCTTCCGCTGATGATGACCAACGCTTGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 |||||
 Db 2612 GTTCTGTGGGATACAGCCAAACACTTGGACTCTTAACCAAGCCCT 2657
 RESULT 14
 US-10-137-865-71
 ; Sequence 71, Application US/10137865
 ; Publication No. US20030032155A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C154
 ; CURRENT APPLICATION NUMBER: US/10-137-865
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 71
 ; LENGTH: 3732
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; US-10-137-865-71
 Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
US-09-978-385-2_COPY_19_738 (1-720) x US-10-137-865-71 (1-3732)	3720.00	83.88%	83.76%	96.15%	3732	717	1	2	138	1
1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaLysP					20					
94 TCACACATTCAGAGAACAGCCCAAGACATTTTGGACAAAGTTTAAACACAGAACCCGAGAC					153					
21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheTyrAsnThrAsnIleThrLysLys					40					
154 CTGTTTATCAAAAGTTCACTGCTTCTTGAATTAACCAAAATTAATCAAGAT					213					
41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr					60					
214 GTCCAAACATGATATATGCTGGGCAAAATGCTGCTTTTAAAGCAACATCCACA					273					
61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln					80					
274 CTGGCCAAATGATATCCACATCAAGAAATTCAGATCTCAGTCAAGCTCAGCTGCAG					333					
81 AlaLeuGlnGlnAsnGlySerSerValLysSerGluAspLysSerLysArgLeuAsnThr					100					
334 GCTCTTCAGCAAAATGGCTTCACAGCTCCTCAGAAAGACAAAGCAAGCTTGAACACA					393					
101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAsnPro					120					
394 ATTCTAAATCAATGAGCCACCATCTACAGTACTGGAAAGTTTGTAAACCAATATCA					453					
121 GlnGluCysLeuLeuLeuGluProGlyLysAsnGluIleMetAlaAsnSerLeuAspTyr					140					
454 CAAGATGCTTATATCTTGAACCAAGTTGAATGAATATGCAAAACAGTTTAACTAC					513					
141 AsnGluArgLeuTrpAlaTrpGluSerTyrPheSerGluValGlyLysGlnLeuArgPro					160					
514 AATGAGAGCGCTGGGCTGGGAAAGCTGAGATGAGAGCGGCAAGACGTCAGGCCA					573					
161 LeuTyrGluGluTyrValValLysLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp					180					
574 TTATATGAGAGATGATGCTGCTTGAATAATGAGATGCAAGCAATATCATATGAGAC					633					
181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspLysTyrAspTyrSer					200					
634 TATGGGATTTATGAGAGAGACATGAAATGAGGATGATGCTATGCTACACG					693					
201 ArgGlyGlnLeuIleGluAspValGlnHisThrPheGluGluIleLysProLeuTyrGlu					220					
694 CGCGGCCAGTTGATGGAAGATGGAACATACCTTGAAGAGATTAACCATTAATATGAA					753					
221 HisLeuHisAlaTyrValAlaGlyAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro					240					
754 CATCTTATGCTATGAGAGGCAAGTATGATGAAGCCATCTCTATATCAATCA					813					
241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeu					260					
814 ATTGATGATCTCCCTGCTCATTTGCTGATGATGAGGATGATTTGAGCAAAATCTG					873					
261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal					280					
874 TACTCTTGAAGCTCTCTTGGACAGAAACCAACATTAATGCTGATGCAATGAGTG					933					
281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal					300					
934 GACGAGCGCTGGGATGACAGAGATATTCAGAGAGCGAGAGAGTCTTGTATCTGTT					993					
301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn					320					
994 GGTCTTCTATATGACTCAAGGATCTGAGAAATTCATGCTAACGAGCCAGAAAT					1053					

QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgTle 340
 Db 1054 GTTCAGAAAGACAGTCTGCCATCCACAGCTTGGGAGGCGGACCTTTCAGGATC 1113
 QY 341 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisIleGluMetGlyHis 360
 Db 1114 CTATGTGCACAAAGGTGACAAATGGACAGCTTCGACAGCCTCATCATGAGTGGGCGAT 1173
 QY 361 IleGlnTrpAspMetAlaValAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 Db 1174 ATCCAGTATGATATGGCATATGCTGCACAACTTTCTGCTAAGAAATGAGACCTAATGAA 1233
 QY 381 GlyPheHisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 Db 1234 GGATTCAGTAAAGCTGTGGGCAATCATGCTCTTTCGACGCGACCTTAAGCATTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 Db 1294 AATATCCATGTGCTTCTGTCACCCGATTTTTCAGAAAGACAAATGAAACAAATAAATTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 Db 1354 CTGGCTAAACAGCACTCAGATGTTGGGAGCTGCGCATTTACTTACATGTTAGAGAG 1413
 QY 441 TrpArgTrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTrpTrp 460
 Db 1414 TGGAGTGGATGGTCTTTAAAGGGGAATTCACAAAGCCAGTGTATATAAAGTGGGG 1473
 QY 461 GluMetLysArgGluIleValGlyValGluProValProHisAspGluThrTyrCys 480
 Db 1474 GAGAGAAAGCGAGATAGTGGGCTGGGAGACCTGGCCCATGATGAACAAATACGT 1533
 QY 481 AspProLysSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
 Db 1534 GACCCCGATCTCTCTTCATGTTTCTGATATGATCATTCATTCGATTAATACAGAG 1593
 QY 501 ThrLeuTrpGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 Db 1594 ACCCTTTCACCAATTCAGATTTCAGAAAGCACTTTGTCAAGAGCTAAACATGAAAGGCCCT 1653
 QY 521 LeuHisLysCysAspPheLysSerAsnSerThrGluAlaGlyLysLeuPhe----- 537
 Db 1654 CTGCACAAATGTGACATCTCAAACTCAACAGACGTGACAGAAACGTGT- GTAAGAAAT 1712
 QY 537 ----- 537
 Db 1713 ACCTCAAAATGTGACCTCTCCTAGTATTCAGTATTACTCATTTCCATGCCATAGTTTG 1772
 QY 537 ----- 537
 Db 1773 TATTTGATTTCTTGTCTTAAAGAAATTTTATGGCTCAAAATGCTCATTTACAA 1832
 QY 537 ----- 537
 Db 1833 ACCAAACATTTAATTTGTGTGACAGACAGAACTAGACATACAAACATTTGGGTGGCCA 1892
 QY 537 ----- 537
 Db 1893 CCTCTTTTCCCTATACATACAGACCTCTCTCTCTCTGTAATTGGAAAGAAAGAGCG 1952
 QY 537 ----- 537
 Db 1953 GTTTAGGAGTGAATATATCTGTAATATGATCTTTTCTTATCTGCAGAACCAAAATTT 2012
 QY 537 ----- 537
 Db 2013 AGCCAACTCAAGAGAAAGAAACCATAGATGTAATATATATGATCATCTGGAACC 2072
 QY 538 ----- 538
 Db 2073 CCTCAAAAGGCCCTGAACCCCTTTTGTGTGACAAATATGCTGAGGCTTGAAATATCA 2132
 QY 546 GluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgPro 565

Db 2133 GAACCTGGACCCCTAGCATTGAAATAATGTTAGAGAC -AAGAACATGATGTAAGGCCA 2191
 QY 566 LeuLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysSerPhe 585
 Db 2192 CTGGCTCAACTCTTTGAGCCCTTATTTACTGGCTGAAACACAGAAAGAAATTCCTTT 2251
 QY 586 ValGlyTrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSe 605
 Db 2252 GTGGAGAGAGATACCGATGAGTCCATATGCACACCAACCAATCAAGTGAAGATAG 2311
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetLysLeuPhe 625
 Db 2312 CCTAAATCAGCTCTTGGAGATTAAGCATATGAAATGGAACGACAAATGATACCTGTT 2371
 QY 625 eArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetI 645
 Db 2432 TCTTTTGGGAGAGAGATGTCGAGTGTATTTGAACCAAGATCTCCTTAATTT 2491
 QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
 Db 2492 CTTTGTACACGACCTAAATATGCTGTATATCATCTAGAACTGAAGTTGAAAGGC 2551
 QY 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAl 685
 Db 2552 CATCAGATGTCCCGGAGCCGCTATCAATGATGCTTCCGTGAAGAACAGAGCCTTGA 2611
 QY 685 aileArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspSerLeuG 705
 Db 2612 GTTTCGTGGGATACAGCAACCACTTGACCTCTTAACAGCCGCCCT 2657
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2657 GTTTCGTGGGATACAGCAACCACTTGACCTCTTAACAGCCGCCCT 2657

RESULT 15
 US-10-140-474-71
 : Sequence 71, Application US/10140474
 : Publication No. US20030032156A1
 : GENERAL INFORMATION:
 : APPLICANT: Baker, Kevin P.
 : APPLICANT: Beresini, Maureen
 : APPLICANT: Deforge, Laura
 : APPLICANT: Desnoyers, Luc
 : APPLICANT: Filvaroff, Ellen
 : APPLICANT: Gao, Wei-Qiang
 : APPLICANT: Gerltsen, Mary E.
 : APPLICANT: Goddard, Audrey
 : APPLICANT: Godowski, Paul J.
 : APPLICANT: Gurney, Austin L.
 : APPLICANT: Sherwood, Steven
 : APPLICANT: Smith, Victoria
 : APPLICANT: Stewart, Timothy A.
 : APPLICANT: Tumas, Daniel
 : APPLICANT: Watanabe, Colin K
 : APPLICANT: Wood, William
 : APPLICANT: Zhang, Zemin
 : TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 : FILE OF INVENTION: ACIDS ENCODING THE SAME
 : FILE REFERENCE: P3330R1C162
 : CURRENT APPLICATION NUMBER: US/10/140,474
 : PRIOR APPLICATION removed - See Palm or File Wrapper
 : NUMBER OF SEQ ID NOS: 550
 : SEQ ID NO 71
 : LENGTH: 3732
 : TYPE: DNA
 : ORGANISM: Homo Sapien
 : US-10-140-474-71
 Alignment Scores:
 Pred. No.: 0 Length: 3732
 Score: 3720.00 Matches: 717

Percent Similarity: 83.88% Conservative: 1
 Best Local Similarity: 83.76% Mismatches: 2
 Query Match: 96.15% Indels: 138
 Gaps: 1
 DB: 9

US-09-978-385-2_COPY_19_738 (1-720) x US-10-140-474-71 (1-3732)

QY	1	SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp	20	DB	1054	GTTCCAGAAAGCAAGTCCCATCCACAGCTTGGGAGCTGGGGAGGGGCACTTCAGGATC	1113
DB	94	TCCACCATGTGGAGAACGGCCAGACATTTTGGACAAAGTTTAACCCAGCAAGCCGGAAGAC	153	QY	341	LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlnHis	360
QY	21	LeuPheTyrGlnSerSerLeuAlaSerThrPheAsnTyrPheAsnHisGluGlnAsn	40	DB	1114	CTTATGTGCAGAAAGGTGACATGAGCGACTTCCTGAGACAGCTCATCATGAGATGGGGCAT	1173
DB	154	CTGTCTATCAAAAGTCTACTTGTCTTGGAAATTAATACACCAATATTTACTGAAAGAAAT	213	QY	361	11leGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu	380
QY	41	ValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuTyrGlnGlnSerThr	60	DB	1174	ATCCAGTATGATGATGCAATATCTGCACAAACCTTTCTGCTGAAGAAATGGAGCTAATGA	1233
DB	214	GTCCAAAACATGATATATGCTGGGACAAATGGCTCCCTTTTAAAGAAACAGTCCACA	273	QY	381	GlyPheHisGluAlaValGlyGlnLeuMetSerLeuSerAlaAlaThrProLysHisLeu	400
QY	61	LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln	80	DB	1234	GGATTCATGAAGCTGTGGGAAATCATGTCATTTTGCAGCCACACTTAAGCATTA	1293
DB	274	CTTGCCCAATGTATCCACTACAAAGAAATTCAGAAATCTCAGCTCAAGCTTCAGCTGAC	333	QY	401	LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe	420
QY	81	AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr	100	DB	1294	AAATCATGTGGCTTCTGTACCCGATTTTCAAGAAACAAATGAAGAAATTAACCTTC	1353
DB	334	GCTCTCAGCAAAATGGGCTCTCAGTCTCTCAGAAACAAAGCAAAACGGTTGAACACA	393	QY	1354	CTGCTCAAAACAGCACTCAGATTTGTGGAGCTGCGCAATTAATCTTACATGTTAGAGA	1413
QY	101	IleLeuAsnThrMetSerThrIleTyrSerThrGlyValCysAsnProAspAsnPro	120	DB	441	TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysTrpTrp	460
DB	394	ATTCTAATACAAATGACACCACTCTACAGTACAGTACGAAAGTTTGTAAACCCAGATTA	453	QY	1414	TGGAGGTGATGGCTTTTAAAGGGAATTTCCAAAGACAGATGATGAAAGAGTGGTG	1473
QY	121	GlnGluCysLeuLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr	140	DB	461	GluMetLysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCys	480
DB	454	CAAAATAGCTTATTACTTGAACACAGTTTGAATGAATGAATGGCAACATTTAGACTAC	513	QY	1474	GAGATACACGAGAGTATTTGGGGTGGAGAACTGTGCCCATGATGAAGAAATACGT	1533
QY	141	AsnGluArgLeuThrPhePheGluSerTrpArgSerGluValGlyLysGlnLeuArgPro	160	DB	481	AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg	500
DB	514	AATGAGAGCTCTGGGCTTGGGAAAGCTGAGATCTGAGTCCGCAAGCCAGCTGAGGCCA	573	QY	1534	GACCCGCAATCTCTGCTTCATGTTCTGATGATTAATCATTCATTCATGATTAACAGAG	1593
QY	161	LeuTyrGlnGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp	180	DB	501	ThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaHisGlnGluPro	520
DB	574	TTATATGAGAGATGTGGCTTGAATAATGAGATGAGCAAGCAAAATCATTTAGAGGAC	633	QY	1594	ACCTTTACCAATTCACATTCAGTTTCAAGAGCACTTTGTCAAGCAGCTTAACAGTAAGAGCCCT	1653
QY	181	TyrGlnAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer	200	DB	521	LeuHisLysCysAspIleSerAsnSerThrGluAlaGlnGlyLysLeuPhe-----	537
DB	634	TATGGGATTTATTTGAGAGAGACTATGAAGTAAATGGGTAGATGCTTATGACTACAGC	693	QY	1654	CTGCACAAATGTGACATCTCAAACTCTACAGAAAGCTGGAGCAAAACTGTT-GTAGAAAT	1712
QY	201	ArgGlyGlnLeuIleGluAspValGlnHisThrPheGlnGluIleLysProLeuTyrGlu	220	DB	537	ACCTCAAAATGTTGAACCTCTCTAGTATTCAGTATTAATCTATTCATTCCTAGCTTGTG	1772
DB	694	CGCGGCCAGTTGATTTGAAGATGTGGAACATATCTTTGAAGAGATTAACCATTTATATGA	753	QY	1773	TATTTATTTCTTTGTTCTAAAAAGAAATTTTATGCTCAAAATGTCCTCATTTACAA	1832
QY	221	HisLeuHisAlaTyrValAlaGlnAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro	240	DB	537	ACCAACATTTAATTTGTGTGTCAGACAGAGAACCTAGACATACAAATTTGGTGGGCA	1892
DB	754	CATCTCATGCTGTATGTGAGGGCAAAATGATGATGATCTCTTCTTATATCATGCA	813	QY	537	CTCTTTTCTCCTATCATTAACACAGCCCTCTCTCTGTATTTGGAAGGAAGAGAGCG	1952
QY	241	IleGlyLysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpHisLeu	260	DB	537	CTTTAGGGTGAATATATCTGTTAATATGATTCATTTCTTTCTTATCTGCCAGAAACAAATTT	2012
DB	814	ATTGGATGCCCTCCCTGCTCATTTGCTTGGATATATGGGGTGAATTTTGGACAAATCTG	873	QY	537	AGCCAACTGTCAGAGAAAGAAACATATGATCATAGATATTAATATGATCATCTGGAAC	2072
QY	261	TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal	280	DB	538	CTTCAAAAGCCCTGAGCCCTTTTGTGTAGCAAAATATCTGAGGCTTGGAAACCA	2132
DB	874	TACCTTTGACAGTCCCTTTTGGACAGAAACCAACATACATGTTACTGATGCAATGGTG	933	QY	546	GluProTrpThrLeuAlaLeuGluAsnValAlaGlyAlaLysAsnMetAsnValArgPro	565
QY	281	AspGlnAlaTrpAspAlaGlnArgIlePheLysGlnAlaGluLysPhePheValSerVal	300	DB	2133	GAACTCTGAGACCTAGCATTTGGAATAATGTTGTAGAGAC-AAAGACATGATATTAAGGCCA	2191
DB	934	GACCAAGCCCTGGATGACAGAGAAATATCAAGAGAGCCGAGAAATCTTTGATATGCTT	993				
QY	301	GlyLeuProAsnMetTrpGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn	320				
DB	994	GGCTTCTTATATGACTCAAGAGATTTGGGAAATTCATGCTTAACGGAGCCAGGAAAT	1053				
QY	321	ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle	340				

QY 566 LeuLeuAsnThrPheGluProLeuPheThrTriPheLeuLysAspGlnAsnLysAsnSerPhe 585
 DB 2192 CTGCTCACTACTGTTGGCCCTTATTTACTGCTGCTGAAGACGAGAACTGTTT 2251
 QY 586 ValGlyTrpSerThrAspTrpSerProTyrValAlaAsp-GlnSerIleLysValArgIleLeu 605
 DB 2252 GTGGAGTGGAGTACCGAGCTGAGTCCATATGACAGCCAAAGCATCAAGTGGATTAAG 2311
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPasnAspAsnGluMetTyrLeuPhe 625
 DB 2312 CCTAAATCAGCTCTTGAGATTAAGCATATGATGAGAACGACAAATGAATGATCCTGTT 2371
 QY 625 eaGserSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetIle 645
 DB 2372 CCATATCTGTGCTATATGCTATGAGGAGTACTTTTAAAGTAAATCAAGATGAT 2431
 QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
 DB 2432 TCTTTTGGGAGAGAGATGTGGAGTGGCTATTTGAAACCAAGATCTCCTTATTT 2491
 QY 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAl 685
 DB 2492 CTTTGCTACTGCACTTAAATGTCTGATATCATTCCTAGAACTGAAGTTGAAAGAGC 2551
 QY 685 aileArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGlu 705
 DB 2552 CATCAGGATGTCCCGAGGCGTATCATATGCTTTCCTGTAAGAACAGACGCTAGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2612 GTTCTGGGAGATACAGCAACACTTGACCTCTATACAGACCCCTC 2657
 RESULT 16
 US-10-142-431-71
 ; Sequence 71, Application US/10142431
 ; Publication No. US20030036179A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Geritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P330RJC251
 ; CURRENT APPLICATION NUMBER: US/10/142,431
 ; PRIOR FILING DATE: 2002-05-10
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 71
 ; LENGTH: 3732
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-142-431-71
 Alignment Scores:
 Pired. No.: 0 Length: 3732
 Score: 3720.00 Matches: 717
 Percent Similarity: 83.88 Conservative: 1
 Best Local Similarity: 83.76 Mismatches: 2

Query Match: 96.15% Indels: 138
 DB: 9 Gaps: 1
 US-09-978-385-2_copy_19_738 (1-720) x US-10-142-431-71 (1-3732)
 QY 1 SerTrpIleGluGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 94 TCCACCATGAGGAGGACGACCAAGCATTTTGGCAAGTTTAAACGAGGAGCCGAGAGAC 153
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPasnTyrAsnThrPasnIleThrGluGluAsn 40
 DB 154 CTGTTCTATCAAGTCACTGCTTCTGGAATTAATACCAATATTACTGAAAGAAAT 213
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerAlaPheLeuLysGluGlnSerThr 60
 DB 214 GTCCAAACATCAATTAATGCTGGGACAAATGCTGCTTTTAAAGAAACACTCCACA 273
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 274 CTGGCCAAATGATTCACACTACAAAGAAATTAAGATTCACAGTCAAGCTTCACCTCAG 333
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 334 GCTTTCAGCAAAATGGGTCTTCAGTGTCTCAGAAAGACAAAGCAACGGTTGACACA 393
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 DB 394 ATTCATTAATACAAAGAGACACATCTACAGTCTGAAAGTTTGTAAACCCAGATTAATCA 453
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 454 CAGAAATCTTATTACTTGAACAGGAGTTGATGAATTAATGCAACAGTTTAACTTAC 513
 QY 141 AsnGluArgLeuTrpAlaTrpGluSerTyrPargSerGluValGlyLysGlnLeuArgPro 160
 DB 514 AATGAGAGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGTGGCGCAAGCAGCTGAGGCCA 573
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 574 TTATATGAAAGATATGTCCTTGAATAATGATGATGATGATGATGATGATGATGATGAT 633
 QY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 DB 634 TATGGGATTAATTTGAGAGGAGTGAAGTGAATGAATGGATGATGATGATGATGATGATGAT 693
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 DB 694 CGCGCGCGTGAATGAAGATGGAACATATCTTGAAGAGTTAAACCATTTATATGA 753
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 754 CATCTCATGCTATGAGGAGGCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 813
 QY 241 IleGlyCysLeuProAlaHisIleLeuGluLysMetTrpGlyArgPheTrpPheTrpAsnLeu 260
 DB 814 ATTGATGCTCTCTGCT 873
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 874 TACTCTTGAACAGTCT 933
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 934 GACCAGGCTGGGATGACAGAGAAATATCAAGAGAGCCGAGAAATCTTTGATCTGTT 993
 QY 301 GlyLeuProAsnMetIleGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 994 GGTCTTCTATATGACTCAAGGATCTGGGAAATTTCAATGCTTAAACGAGCCAGCAAAAT 1053
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 DB 1054 GTTCAGAAACAGTCTGCTCATCCACAGCTTGGGACCTGGGGAAGGCGACCTTCAGAGATC 1113

```

QY 341 LeuMetCysThrIysValThrMetAspAspPheLeuThrAlaHisIleGluMetGlyHis 360
DB 1114 CTATATGTCACAAAGGAGACATGGACACTCTCTGACAGCTCTCTGAGATGGGCGAT 1173
QY 361 IleGlnIyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
DB 1174 ATCCAGATATGATATGGATATATGTCACAAACCTTTTCTGCTAAGAAATGAGAGCTAATGAA 1233
QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
DB 1234 GGATTCATGATGAGCTGTGGGAAATCATGCTCTCTGACAGCCACCACTAAGCAATTTA 1293
QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
DB 1294 AAATTCATGCTGTCTGTCTGTCACCGGATTTTCAGAAACATGAAACAGAAATTAACCTTC 1353
QY 421 LeuLeuIysGlnAlaLeuThrIleValGlyThrLeuProPheThrIleMetLeuGluLys 440
DB 1354 CTGCTCAAAACAGACACGACAGATGTTGGGACCTCTGCCATTTTACTTACATGTTAGAGAG 1413
QY 441 ThrAlaGlyPheMetValPheLysGlyGluIleProLysAspGlnThrPheLysLysTrp 460
DB 1414 TGGAGGTGGATGCTCTTTAAAGGGGAAATTCACCAAGACAGCTGATGAAAGAAAGTGGTG 1473
QY 461 GluMetLysArgGluIleValGlyValGluProValProHisAspGluThrIleCys 480
DB 1474 GAGATGAGGAGGAGATATGTTGGGGGTGGGACCTGTGCCCGGATGAAACATTAATCTGT 1533
QY 481 AspProLysSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
DB 1534 GACCCCGCATCTGTTCCATGTTTCTGATGATATACATTCATTCATTCATTCATTCATTCAT 1593
QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
DB 1594 ACCCTTACCAATTCCTGATTCAGAAAGACCTTGTCTGACAGCTTAAACATGAAAGGCCCT 1653
QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysPhe 537
DB 1654 CTCGACAAATGTGACATCTCAAACTCTACAGAACCTGACAGAACTGTT-CTAAGAAAT 1712
QY 537 537
DB 1713 ACCTCAAAATGTGACACCTCTCTAGTATTCATTCATTCATTCATTCATTCATTCATTCAT 1772
QY 537 537
DB 1773 TATTTGATTTCTTTGTTCTAATAAAGAAATTTATGGCCCAAAATGCTCATTTTACAA 1832
QY 537 537
DB 1833 ACCAAACATTTAATTTGTGTGACAGACAGAACTACACCATCAACATTTGGGTGGCCCA 1892
QY 537 537
DB 1893 CCTCTTTCTCCATCATTAACCTACAGCCCTCTCTCTGCTATTTGGAAGGAAAGAGCG 1952
QY 537 537
DB 1953 GTTTAGGTGGAATATATCTGTATATATGATCTCTTTCTTATCTGCCAGAGCAAAATTT 2012
QY 537 537
DB 2013 AGCCAAATCAAGAGAGAAACCATAGATCATAGATGAATATATGATCATCTGGAACC 2072
QY 538 538
DB 2073 CCTCAAAAGGCCCTGAACCCCTTTTGTGTAGCATATATGCTGAGGCTTGGAATATCA 2132
QY 546 GluProTyrPheLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgPro 565
DB 2133 GAACCTTGAGCCCTTGATGATGGAATATGTTGTAGAGC-AAGAACATGAATGTAAGGCCA 2191
QY 566 LeuLeuAsnTyrPheGluProLeuPheThrIlePheLysAspGlnAsnLysAsnSerPhe 585

```

```

DB 2192 CTGCTCACTACTTATGAGCCCTTATTTACCTGCTGAAAGACCAACAGATCTTTT 2251
QY 586 ValGlyTrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSe 605
DB 2252 GTGGATGAGAGTACCGACATGAGATCCATATGACAGACCAACATCAAGATGAGATGAG 2311
QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPheAsnAspAsnGluMetTyrLeuP 625
DB 2312 CCTTAAATACGCTCTGGAGATTAACCATATGAAATGGAACGACATGAATATGATCTGTT 2371
QY 625 eArgSerSerValAlaTyrAlaMetArgGlnTyrPheLysValLysAsnGluMetI 645
DB 2372 CCGATCATCTGTTGCTATATGCTATGAGGACAGTACTTTTAAAGTAAATCAATGAT 2431
QY 2432 TCTTTTGGGAGAGAGATGTCGAGTGCATATTTGAACCAAGATCTCTTAAATTT 2491
DB 2492 CTCTTGCCTGACACCTTAAATGCTGTATATCATCTTCAACTGAAAGTGAAGAGC 2551
QY 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAl 685
DB 2552 CATCAGAGATGTCGGGAGCCGATTCATATGATGCTTCCCTGATGACACAGCCTAGA 2611
QY 705 upheLeuGlyIleGlnProThrLeuGluIleProAsnGlnProPro 720
DB 2612 GTTCTGGGATGATACAGCAACCACTTGACCTCTTACAGACCCCT 2657

RESULT 17
US-10-143-114-71
; Sequence 71, Application US/10143114
; Publication No. US20030036180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gudowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C211
; CURRENT APPLICATION NUMBER: US/10/143,114
; PRIOR APPLICATION DATE: 2002-05-09
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 71
; LENGTH: 3732
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-143-114-71

Alignment Scores:
Pred. No.: 0
Score: 3720.00
Percent Similarity: 83.88%
Best local Similarity: 83.76%
Query Match: 96.15%
DB: 9
Length: 3732
Matches: 717
Conservative: 1
Mismatch: 2
Indels: 138
Gaps: 1

```


OY 586 ValGlyTyrSerThrAspTrpSerProTyrAlaAsp-GlnSerIleuValArgIleSe 605
 DB 2252 GTGGAGTGGAGTACCCAGCTGAGTCCATATGACAGCCAAAGCATCAAGTGAAG 2311
 OY 605 rleuValSerAlaLeuGlyAspIleValTyrGluTyrPheAsnAspAsnGluMetIleu 625
 DB 2312 CCTAAATACAGCTTGTGAGATTAAGCATATGATGAGACAAATCAATGATGAT 2371
 OY 625 eaTgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuValIleuValSerIleu 645
 DB 2372 CCGATCATCTGTGATATCTATGAGGAGCTCTTTTAAAGTAAATATCATGAT 2431
 OY 645 elePheGlyGluGluValAspValArgValAlaAsnLeuValProArgIleSerPheAsn 665
 DB 2432 TCTTTTGGGAGAGAGATGCGAGTGGTAAATTTCAACCAAGATCTCTTATTT 2491
 OY 665 ePheValThrAlaProIleAsnValSerAspIleIleProArgThrGluValGluVal 685
 DB 2492 CTTGTCTACCTGACCTAAATATGCTGTGATATCATCTTGAAGTGAAGTGAAGGC 2551
 OY 685 alleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeu 705
 DB 2552 CATCAGGATGTCCTGGAGAGCCCTATCATATGATGCTTCCGCTGATGATCAACACCTTGA 2611
 OY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2612 GTTCTGGGATACAGCCACACCTGGAGCTCTTGAACACAGCCCT 2657

RESULT 18
 US-10-140-002-71
 : Sequence 71, Application US/10140002
 : Publication No. US20030037623A1
 : GENERAL INFORMATION:

: APPLICANT: Baker, Kevin P.
 : APPLICANT: Beresini, Maureen
 : APPLICANT: DeForge, Laura
 : APPLICANT: Desnoyers, Luc
 : APPLICANT: Filvaroff, Ellen
 : APPLICANT: Gao, Wei-Qiang
 : APPLICANT: Gerritsen, Mary E.
 : APPLICANT: Goddard, Audrey
 : APPLICANT: Godowski, Paul J.
 : APPLICANT: Gurney, Austin L.
 : APPLICANT: Sherwood, Steven
 : APPLICANT: Smith, Victoria
 : APPLICANT: Stewart, Timothy A.
 : APPLICANT: Tumas, Daniel
 : APPLICANT: Watanabe, Colin K
 : APPLICANT: Wood, William
 : APPLICANT: Zhang, Zemin
 : TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 : FILE REFERENCE: P333081C59
 : CURRENT APPLICATION NUMBER: US/10/140,002
 : PRIORITY FILING DATE: 2002-05-06
 : PRIOR Application removed - See Palm or File Wrapper
 : NUMBER OF SEQ ID NOS: 550
 : SEQ ID NO 71
 : LENGTH: 3732
 : TYPE: DNA
 : ORGANISM: Homo Sapien
 : US-10-140-002-71

Alignment Scores:
 Pred. No.: 0 Length: 3732
 Score: 3720.00 Matches: 717
 Percent Similarity: 83.88% Conservative: 1
 Best Local Similarity: 83.76% Mismatches: 2
 Query Match: 96.15% Indels: 138
 DB: 9 Gaps: 1

US-09-978-385-2_COPY_19_738 (1-720) x US-10-140-002-71 (1-3732)

OY 1 SerThrIleGluGluAlaValLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 94 TCCACCATGAGAGAACAGCCAAAGCATTTTGGACAAAGTTAAACAGGAAGCCGAAGAC 153
 OY 21 LeuPheTyrGlnSerSerLeuAlaSerThrPheAsnTyrAsnThrAsnIleThrGluGluAsn 40
 DB 154 GCTTCTTATCAAAAGTTCCTCTTCTTGGATTTAAACACCAATATTAATGACAGAT 213
 OY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerAlaPheLeuValGluGlnSerThr 60
 DB 214 GTCCAAATATGATATATCTGGGACAAATGGTCCCTTTTAAAGGAACAGTCCACA 273
 OY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnIleGln 80
 DB 274 CTGGCCAAATGATCATCTACAGAAATTCAGAAATTCACAGTCAAGCTTCACCTGAG 333
 OY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 334 GCTCTTCAGCAAAATGGGCTTCAGTCTCTCAGAAAGACAAAGCAAGCGTTGAACACA 393
 OY 394 ATTCTAAATACATGAGACACCATCTACAGTACGAAAGTTGTAAACCATGATATCA 453
 OY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 454 CAAGAAATGCTTATTAATCTTGAACAGGTTGAATGAATATATGCAAAAGTTAGACTAC 513
 OY 141 AsnGluArgLeuThrAlaTrpGluSerTyrParSerGlyValGlyGlnLeuAspPro 160
 DB 514 ATTCAGAGCTCTGGGCTTGGAAAGCTGAGATCTGAGAGTGGCAAGCAGCGAGCA 573
 OY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 574 TTATATGAAGATGATGCTCTGAAATATGAGATGCAAGCAATATATATGAGAGAC 633
 OY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 DB 634 TATGGGATTAATGGAGAGGAGACTGAAAGTAAATGGGTATGATGCTATGACTACAGC 693
 OY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 DB 694 CGCGCCAGTTGANTGAAAGATGGAACATACCTTTGAAGAGATTAACCATTAATATGAA 753
 OY 221 HisLeuHisAlaTyrValAlaGalaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 754 CATCTCATGCTCTATGAGGCAAGTGTATGATGATGATGATGATGATGATGATGATGAT 813
 OY 241 IleGlyCysLeuProAlaHisIleLeuGlyAsnMetTrpGlyArgPheTrpThrAsnLeu 260
 DB 814 ATTGATGCTCTCTGCTCATTTGCTGGATATGATGGGTATGATTTTGGACAAATCTG 873
 OY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 874 TACTCTTGAAGATTCCTCTTGGACAGAAACCAATAGATGATGATGATGATGATGATGAT 933
 OY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 934 GACCAGGCTGGGATGACAGAGATATTCAGAGGAGCGAGAGTCTTGTATCTGTT 993
 OY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 994 GCTCTTCTTATATGATCTCAAGATCTGGGAAATTCATGCTAAAGGAGCCAGGAAT 1053
 OY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 DB 1054 GTTCAGAAAGCAGCTGCGCATCCACAGCTTGGAGCCTGGGAGAGCGGACTTCAGGATC 1113
 OY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHis 360
 DB 1114 CTTATATGACAAAGAGTGACATGACAGCTTCTGACAGCTCATATGATGATGGGCAAT 1173

Db	2252	GTGGGATGGAGTACCGAGCTGAGTCCATATGCAAGCCAAAGCATCAAAAGTGGAGTATAC	23111
QY	605	rleuIySSerAlaIeugIyAspIyAlaIeIyGluTTPasnaSpasngIumetYrLeuph	625
Db	2312	CCTAAATATCAGCTCTTTGGAGATAAAGCAATATCAAAAGAACACACATATGAAAGTCACTGTT	2371Y
QY	635	eArGSeSerValAlaIeYrAlaMetArGInIeYrPheLeuIySaIlySaSangImetIi	645
Db	2372	CCGATCATCTGTTCGATATTCGATATGAGGACAGTACTTTTAAAGTAAAAATTCAGATCAT	24311
QY	645	eIeuhheGlyGluGuaSpvAlArGvAlAlaIsnLeuIySProArGIIeSerPheasnph	665
Db	2432	TCTTTTGGGGGGAGGAGATGTGCGAGTGGCTAAATTGAAACCAAGATCTCTTTAAATTT	2491Y
QY	665	ePheValrThrAlaProIySaSnValSerAspIleIeProArGThrGluValGluYsaI	685
Db	2492	CTTTGTCACTGCACCTTAAAAATGTGTGATATCATTTCTTGAACCTGAAGTTGAAAAAGGC	2551Y
QY	685	aIIeArGmetSerArGSeSerArGIIeAsnAspAlaPheArGLeuAsnAspAsnSerLeuGI	705
Db	2552	CATCAGGAGTGTCCCGGAGCCGCTATCAAAAGATGCTTTCCGTGTGATGACAAACACCTTGA	2611Y
QY	705	uPheLeuGlyIleGInProThrLeuGlyProProAsnGInProPro 720	
Db	2612	GTTTCTGGGGATACAGCCAAACACTTGGAGCTCCCTTAAACAGCCCTT 2657	

QY 1 serThrIleGluGluGln

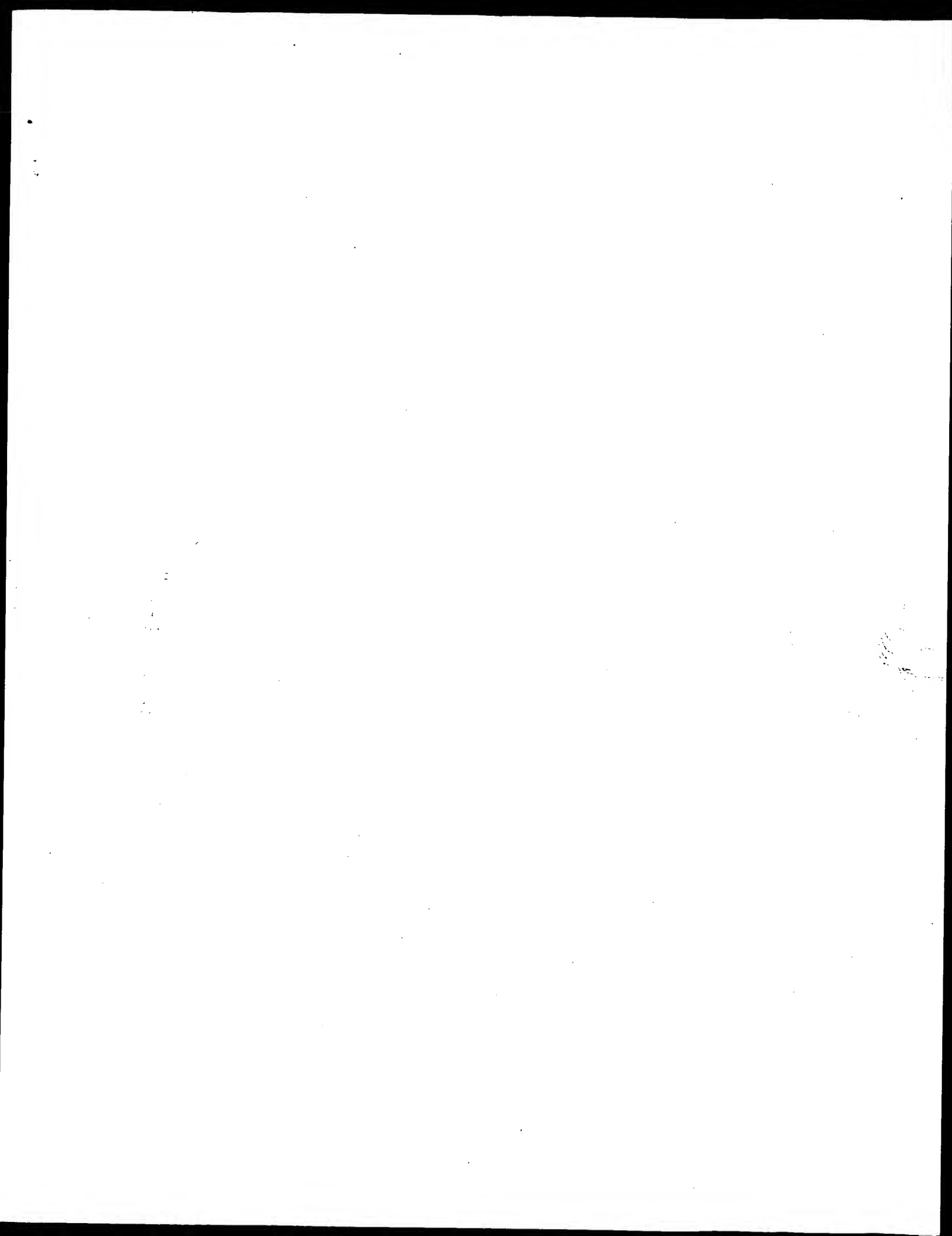
US-09-978-385-2_COPY_19_738 (1-720) x US-10-142-419-71 (1-3732)

Db	1174	ATCCAGATGATATGAGCATATGCTGCACAAACCTTTCTTGCTAAGAAATGAGCAATATGAA	1233
QY	381	GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaIleThrProLysHisLeu	400
Db	1234	GGATTCATCAAGCTGTGGGGAATCATGTCTCTCTCAGCCACACCTAAGCATTTA	1293
QY	401	LysSerIleGlyLeuSerProAspPheGlnGluAspGlnuThrGluIleAsnPhe	420
Db	1294	AAATCCATGTGCTCTTGTCACCCGATTTCAGAAACAAATGAAACAAATTAACCTTC	1353
QY	421	LeuLeuYsgIlnalaleuthrIleValaGlyThrLeuProPheThrTrpMetLeuGluLys	440
Db	1354	CTGCTCAACAGACACACACAGATGTGTGGGAGCTGTGCCATTTCCTACATGTTAAGACAAG	1413
QY	441	TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysTrpTrp	460
Db	1414	TGGAGTGGATGGCTTTTAAAGGGGAATTCCTCCAAAGACAGATGGATGAAGAAAGTGGTGG	1473
QY	461	GluMetLysArgGluIleValaGlyValValGlnProAlaProHisAspGluThrTrpCys	480
Db	1474	GAGATGGAAGCAGAGATAGTGGGGGTGGAGACCTGTGCCCATGATGAAACATCTGTGT	1533
QY	481	AspProAlaSerLeuPheHisValSerAsnAspTrpSerPheIleArgTrpThrArg	500
Db	1534	GACCCCGCATCTGCTTCATCATGTTCGAGATACATCATTCATTCATCATATATACACAGG	1593
QY	501	ThrLeuTrpGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGlyLysPro	520
Db	1594	ACCCTTACCAATTCAGATTCAGAAAGCACTTTGTCAAGCAGCTAAACATGAAGGCCCT	1653
QY	521	LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyGlnLysLeuPhe	537
Db	1654	CTGCACAAATGTGACATCTCAAACTCTACAGAAAGCTGACAGAAACGTGT- GTAAGAAAT	1712
QY	537	-----	537
Db	1713	ACCTCAAAATGTGAACCTCTCCAGTATTCAGTATTAATACATTTCCATGGCTAGCTTG	1772
QY	537	-----	537
Db	1773	TATTTGATTTCTTTGTTCTATAAAGAAATTTATGGCCTCAAAATGTCCTATTATCAA	1832
QY	537	-----	537
Db	1833	ACCAAAACATTTAATTTGTGTGTCAGACAGAAACCTAGAACATACAAATTTGGTGGGCCA	1892
QY	537	-----	537
Db	1893	CCTCTTTTCTCCCATCATTAACATACAGCCCTCTCTCTCGTAAATGGAAAGAAAGACG	1952
QY	537	-----	537
Db	1953	GTTTAGGGTGAATATATCTGTAAATATGATCATCTTTCTTATCGCCAGAAACAATTT	2012
QY	537	-----	537
Db	2013	AGCCAAGTCAAAGAGAGAAACCATAGATCATAGATGTAATATATGATCATCTGGAACC	2072
QY	538	-----	545
Db	2073	CCTCAAAAGGCCCTGCAACCCCTTTTGTGTGTAGCAATATGCTGAGGCTTGGAAAATCA	2132
QY	546	GluProTrpThrLeuAlaLeuGlnAsnValValGlyAlaLysAsnMetAsnValArgPro	565
Db	2133	GAACCCCTGAGCCCTAGCATTTGGAATAATGTTGAGAGC- AAGAACATGATGTAAAGCCA	2191
QY	566	LeuLeuAsnTrpPheGlnProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe	585
Db	2192	CTGCTCAACTACTTTGAGCCCTTATTATCTCGCGGAAGACCAACAAACAAATTCCTTT	2251
QY	586	ValGlyTrpSerThrAspTrpSerProLysAlaAsp- GlnSerIleLysValArgLysLe	605
Db	2232	GTGGGATGAGAGTACCGACTGGAGTCCATATGACAGCCCAAGACATCAATGATGAGATATAG	2311

QY 381 GlyPheHisGluAlaValGlyIleMetSerLeuSerAlaAlaThrProIleSHSLeu 400
 DB 1234 GGATTCATGAAGCTCTGGGAAATCATGCTCTTTCGACCCACACCTTAAGCTTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsn 420
 DB 1294 AATTCATGCTGCTTGTGACCCGATTTTCAGAAAGACATGAAACAGAAATTAACCTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrIleLeuGluLys 440
 DB 1354 CTGCTCAAAACACACCTGAGATGTTGGACCTCCCATTTACTTACATGTTAGGAAG 1413
 QY 441 TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrp 460
 DB 1414 TCGAGGTGATGCTCTTTAAAGGGAATCCCAAGACACGATGAAAGAGTGTGG 1473
 QY 461 GluMetLysArgGluIleValGlyValValGluProValProHisAspGluThrIle 480
 DB 1474 GAGATGAAGCGAGATAGTGGGGTGGTGAACCTGTCCTGATGAAACATGCTGT 1533
 QY 481 AspProLysSerLeuPheHisValSerAsnAspLysSerPheIleArgTrpThrArg 500
 DB 1534 GACCCGCACTCTGCTTCATGTTCTGATGATTACTCATTCATTGATTTACACAGG 1593
 QY 501 ThrLeuTrpGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaAlaLysSHSLeuGlyPro 520
 DB 1594 ACCCTTACCATTCGATTTCAAGAGCACTTTGTCAACGACTTAACGAGGCCCT 1653
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe 537
 DB 1654 CTGCACAAATGTGACATCTCAAACTACAGAAAGCTGACAGAAACGTGT- GTAAGAAT 1712
 QY 537 537
 DB 1713 ACCTCAAATGTGAACCTCTCAGATTTCAGTATTACTCATTTCCATGCTAGGTTG 1772
 QY 537 537
 DB 1773 TATTTGATTTCTGTTCTTAAAGAAATTTTATGGCTCAAAATGCTCATTTACAA 1832
 QY 537 537
 DB 1833 ACCAAACATTAATTTGTGTGACAGAGAAACCTAGACCAATCAACAATTTGGTGGCCA 1892
 QY 537 537
 DB 1893 CCTCTTTCTCCATCATTAATAACAGCCCTCTCTCTGTTAAATTGAGAGAAAGACG 1952
 QY 537 537
 DB 1953 GTTTAGGCTGAAATATATCTGTTAATGATCTTTCTTATCTGCGAAGCAAAATTT 2012
 QY 537 537
 DB 2013 AGCCAAGTCAAGAGAAACCATAGATCATAGATTAATATATGTCATCTGGAAC 2072
 QY 538 538
 DB 2073 CCTCAAAAGGCCCTGAAACCCCTTTTGTGTACCAATATGCTGAGGCTTGGAATAATCA 2132
 QY 546 GluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgPro 565
 DB 2133 GAACCTTGACCTTGACATTTGAAATGTTGAGAGC- AAGACATGTAATGTAAGGCCA 2191
 QY 566 LeuLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 585
 DB 2192 CTGCTCAACTACTTGAAGCCCTTATTTACTGCTGAAAGACACAGAAATCTCTTT 2251
 QY 586 ValGlyTrpSerThrAspTrpSerProIleAlaAsp- GlnSerIleLysValArgIle 605
 DB 2252 GTGGGATGAGATACCGATGAGATGATGACAGCCCAAGCATCAAGTGAAGATTAAG 2311
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTrpGluTrpAsnAspAsnGluMetLysLeuPhe 625

DB 2312 CCTAAATCAGCTCTTGAGATTAAGCATATGAATGGAACGACATGAAATGTAACCTGTT 2371
 QY 625 eArgSerSerValAlaTrpAlaMetArgGlnTrpPheLeuLysValLysAsnGlnMetI 645
 DB 2372 CCGATCATCTGTGATATGCTATGAGCAGTACTTTTAAAGTAAATAATCAGATGAT 2431
 QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
 DB 2432 TCTTTTGGGAGAGAGATGTCGAGTGAATTTGAACCAAGAAATCTCTTAAATTT 2491
 QY 665 ePheValThrAlaProLysAsnValSerAspIleLeuProArgTrpGluValGluLysAl 685
 DB 2492 CTTTGCTACTGCACCTAAATATGTCTGATATCATTTCTTGAAGCTGAAAGTTGAAAGGC 2551
 QY 685 aIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuG 705
 DB 2552 CATCAGAGATGCTCCGAGGCGTATCATATGATGCTTCGTTGATGACCAACAGCCCTAGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2612 GTTCTGGGAGATACAGCCAAACACTTGACCTCTTAACCAAGCCCT 2657

Search completed: May 26, 2003, 20:28:02
 Job time : 349 secs



Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

11
12
13
14
15
16
17
18
19
20
21
22

AL551235 AL

BG962298 6

BT561069 6

DO122010
AA162058 MS

AAU006J30 6
PT561359 6

BE 169139 6
PC438050 6

BM030353 4
PD657069 1

B18264/1 6
DTA1E007 CBM560537 *A*

BG772384 E

AL039738 DK

BI196733 6

BI686319 6

BE784147 6

BI914635 6

AW260204

BI825221

BI103238

BM290186

BF783379

AT971658 W.

BM000000
BM000001
BM000002

B0947556

near HIC 1
(post-1 day)

brata; Eutel

JOURNAL

Submitted (14-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcgaps-remail.nih.gov

COMMENT

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
contact: amandahsystemsbiology.org
Anup Madan, Jessica Fahey, Erin Heltom, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.gov>
Series: IRAC Plate: 34 Row: 1 Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 11225608
This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers

1..3337
/organism="Homo sapiens"
/db_xref="locusID:59272"
/db_xref="taxon:9606"
/clone="IMAGE:4830668"
/tissue_type="Testis"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"

BASE COUNT 1023 a 634 c 756 g 924 t
ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3337
Score: 3827.00 Matches: 717
Percent Similarity: 99.728 Conservative: 1
Best Local Similarity: 99.588 Mismatches: 2
Query Match: 98.918 Indels: 1
DB: 11 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x BC032938 (1-3337)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
94 TCCACCATGAGGAAACAGGCCAAGACATTTTGGACAAAGTTTAAACACGAGCCGAGAC 153
QY 21 LeuPheTYGlnSerSerLeuAlaSerTPAsnTYrAsnThrAsnIleThhGluGluAsn 40
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
154 CTGTTCTATCAAGAGTTCACCTGCTCTTGGAATTTAAACACCAATATTACTGAAAGAGAT 213
QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTTPSerAlaPheLeuLysGluGlnSerThr 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
214 GTCCAAAACATCAATTAATGCTGGGAGCAATGCTGCTTTTAAAGAGAACAGTCCCA 273
QY 61 LeuAlaGlnMetTYrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeu 80
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
274 CTGGCCCAATGATATCCACTACAGAAATTCAGAAATTCACAGTCACAGTCCAGCTGAG 333
QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysATGLeuAsnThr 100
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
334 GCTCTTCAGCAAAATGGGTCTTCAGTGTCTCAGACAGACCAAGACCAAGGTTTAAACCA 393
QY 101 IleLeuAsnThrMetSerThrIleTYrSerThrGlyLysValCysAsnProAspAsnPro 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
394 ATTCATAATACAAATGACACCATCTACAGTACTGAAAAAGTTTAAACCCAGATTAATCCA 453
QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTYr 140
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
454 CAAAGATGCTTATTACTTGAACAGAGTTGAAATGAATTAATGCAACAGATTGACTAC 513

QY 141 AsnGluArgLeuThrPalaTrpLeuSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
514 AATGAGAGGCTCGGGCTGGGAAAGCTGGAATCTGAGTCGCGCAACAGCTGAGGCCA 573
QY 161 LeuTYrGluGluTYrValValLeuLysAsnGluMetAlaArgAlaAsnHisTYrGluAsp 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
574 TTATATGAAGATATGATGCTGCTTGAATAATGATGCGCAAGCAATCATATGAGAGC 633
QY 181 TYrGlyAspTYrTrpArgGlyAspTYrGluValAsnGlyValAspGlyTYrAspTYrSer 200
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
634 TATGGGAGTATTGAGAGAGAGACTATGAGTAATGGGATAGATGCTATGACTACAGC 693
QY 201 ATGGGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTYrGlu 220
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
694 CCGCGCCAGTTCATTTGAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 753
QY 221 HisLeuHisAlaTYrValAlaArgAlaLysLeuMetAsnAlaTYrProSerTYrIleSerPro 240
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
754 CATCTTCATGCTATGATGAGGCAAGTTGATGTAATGCTTCTCTATATATGATCA 813
QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpAsnLeu 260
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
814 ATGGATGCTCCCTGCTCATTTGCTGCTATGATGCTGCTGCTGCTGCTGCTGCTGCT 873
QY 261 TYrSerLeuThrValProPheGlyGluLysProAsnIleAspValThrAspAlaMetVal 280
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
874 TACTCTTTCAGACAGTCCCTTTGGACAGAAACCAATGATGCTGATGCTGATGCTG 933
QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
934 GACCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 992
QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
993 GCTCTTCTTATATGACCTCAAGATCTGGAATCTGGAATCTGGAATCTGGAATCTG 1052
QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuLysGlyAspPheArgIle 340
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1053 GTTCAGAAAGAGCTGCTATGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1112
QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHis 360
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1113 CTATATGTCACAAAGGAGCAATGACAGCTTCTGACACCTCATCATGATGAGTGGCAT 1172
QY 361 IleGlnTYrAspMetAlaTYrAlaAlaGlnProPheLeuLeuArgAsnGluAlaAsnGlu 380
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1173 ATCCAGATATGATATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1232
QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1233 GGAATTCATGAAAGCTGCTGGGAAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1292
QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPro 420
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1293 AATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1352
QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTYrMetLeuGlyLys 440
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1353 CTGCTCAAAACACACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1412
QY 441 TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrp 460
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1413 TGGAGGTGGATGCTTAAAGGGGAAATTCACCAAGACAGTGTGTAAGAAAAAGTGTG 1472
QY 461 GluMetLysArgGluIleValGlyValValGluProValProHisAspGluThrTYrCys 480
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1473 GATATGAAGAGAGATGATGAGGAGTGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1532
QY 481 AspProAlaSerLeuPheHisValSerAsnAspTYrSerPheIleArgTYrTYrArg 500
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1533 GACCCCGCATCTCTGTTTCATGATGCTATGATGCTATGCTATGCTATGCTATGCTATG 1592

OY		501	ThrlenturTcInPhegInPhegIngluaIaleucCysGlnAlaIalAylstHslugLpyro	520
Db		1593	ACCCTTTACCAATTCACGGTTTCAAGAAGCACCTTGTCACACCGCTAAACCTGAAGGCCCT	1652
OY		521	LenhIslysgYasApIlleSerAsnSerThGlalAgLyGlnIySLeupheasnMetLeu	540
Db		1653	CTGCACAANAgtgcAtCTCAAACTCTACAGAAAGcttgAGAAAcGTtCCAAATANGCTg	1712
OY		541	ArgLeugLIySserGIuProTrpThrLeualaleuGLuAsnValValGIyAlalyAsn	560
Db		1713	AGCGTTGGAAAATCAGAACCCCTGAGACCCCTGACATGGAAAAATGTTGAGAGCAAAAGAAC	1772
OY		561	MethsnValarProlleuleuAsnTyPhagLIuProLeupherhTrTpIeulyAspGln	580
Db		1773	ATGAAATGAAGCCACgCTCtCAACTACTTtGAGCCCTTAATTtTACCTGGCTGCAAAAGCACg	1832
OY		581	AsnIyAsnSerPhevalGIyTrpSerThrAspTrpSerProTyralaaSpGlnSerIle	600
Db		1833	AACAAGAAATCTTTTGGTGGAGTGAAGTACCGACAGCTGAGTCCATATGACAGACAAAGCATC	1892
OY		601	LysValarqIIeserIeulySserAlaleuGLysAspIyAlatyrgLUtTPAsnAspAsn	620
Db		1893	AAAGGAGAGATTAAGCTTAATAATCAGCTTGGAGATTAAGCATATGATGAGACGACAAAT	1952
OY		621	GluMetTyLeupheargSerSerValAlatryAlametarGlnTyrPheLeuLYsVal	640
Db		1953	GAAATGTACGTGTTCCGATCATCTGTTGGCATATGCTATGAGGAGCAGTCTTTTAAAAATA	2012
OY		641	LysAsngInMeTIleLeupheGLyGUlUAspValArtyValAlalaenLeuLYsProArg	660
Db		2013	AAAAATCAGATGATCTTTTGGGAGAGAGATGTGCAGAGGGCTAAATTATAAACACAGA	2072
OY		661	IleserPheasnPhepheValThrAlaProLYsAsnValIserAspIleIleProArgThr	680
Db		2073	ATCTCTTAAATTTCTTTGTCACTGCACCTAAAAATGTGTCGATATCATTTCTTAAGAACT	2132
OY		681	GIuValIGLuLYsAlAlIleargMetSerArgSerArgIleaAsnAspAlaPheargLeuAsn	700
Db		2133	GAGGTGGAAGAGCCATCAGATGATGCCGGACCCGTATCATGATGATGCTTCGCTCGAT	2192
OY		701	AspAsnSerLeuGluPheLeuGLyIleGlnProThrIeugLyProProAsngInProPro	720
Db		2193	GACACACGCTTAGATTTCTGGGGATACAGCCAAACACTTGAGACCTCTTAACAGAGCCCCCT	2252
RESULT 2				
AL551235		987 bp	mRNA	linear EST 16-FEB-2001
LOCUS		AL551235 LFI_NFL006.PL2	Homo sapiens cDNA clone CS0D1041YM14	5
DEFINITION		prime, mRNA sequence.		
ACCESSION		AL551235		
VERSION		AL551235.1	GI:12888986	
KEYWORDS		EST.		
SOURCE		human.		
ORGANISM		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE		1 (bases 1 to 987)		
AUTHORS		Li,W.B., Gruber,C., Jesse,J. and Polayer,D.		
TITLE		Full-length cDNA libraries and normalization		
JOURNAL		Unpublished (2001)		
COMMENT		Contact: Genoscope		
		Genoscope - Centre National de Sequencage		
		Bp 191 91006 EVRY cedex - France		
		Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
		Location/Qualifiers		
FEATURES		1..987		
source		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/clone="CS0D1041YM14"		
		/clone_1lb="LFI_NFL006.PL2"		
		/tissue_type="placenta"		
		/note=Vector: PCWSPORT 6; Site:1: NotI; 1st strand cDNA		
		was primed with a NOTI-oligo(dT) primer. Five prime end		

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 322 a 182 c 211 g 270 t 2 others

ORIGIN

Alignment Scores:

Proed. No.:	3.84e-125	Length:	987
Score:	1112.00	Matches:	211
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	28.74%	Indels:	0
Db:	9	Gaps:	0

US-09-978-385-2_COPY_19_738 (1-720) x AL551235 (1-987)

QY	510	AlaLeuCySGlnAlaAlaAlaYSHisGLuGlyProLeuHisLysCysAspIleSerAsnSer	529
Db	3	GCACCTGTTCTACAGACGCTTAAACATGAAGGCCCTCTGCACAAATGTGACATCTAAACCTG	62
QY	530	ThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProTyrPthr	549
Db	63	ACAGAGCGTGACACGAACCTGTTCAATATGCTGCGCTTGAAATTCAGAACCTCGAC	122
QY	550	LeuAlaLeuGluAsnValAlaGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyr	569
Db	123	CTAGCATTTGGAAATGTTGTATGGAGCAACAGACATGATGTAAAGCCACGCTCAACTAC	182
QY	570	PheGluProLeuPheThrTyrPheLysAspGlnAsnLysAsnSerPheValGlyTyrPser	589
Db	183	TTTGAGCCCTATTACTCGCTGCTAAAGACCAGAACAGAAATCTTTGTGGATGGAGT	242
QY	590	ThrAspTyrPserProTyrAlaAspGlnSerIleLysValArgIleSerLeuLysSerAla	609
Db	243	ACCGACTGGAGTCCATATGACAGACCAGCAACATCAAGTAGAGATTAACCTTAATACGCT	302
QY	610	LeuGlyAspLysAlaTyrGluTyrPAsnAspAsnGlnMetTyrLeuPheArgSerSerVal	629
Db	303	CTTGAGAGTAAAGCATATGATGAAACGACATGAAATGTAACCTTCGATCATCTGTT	362
QY	630	AlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetIleLeuPheGlyGlu	649
Db	363	GCATATGCTATGAGGACAGTACTTTTAAAGTAAATCAAGATGATCTTTTGGGGAG	422
QY	650	GluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePheValThrAla	669
Db	423	GAGGATGTCCGAGTGGCTATTTGGAACCAAGAAATCTCTTAATTTCTTGTGCATGCA	482
QY	670	ProLysAsnValSerAspIleIleProArgThrGluValAlaGlyLysAlaIleArgMetSer	689
Db	483	CCTAATAAATGTCTCGATATCATTTCTCTGAACTGAGTTGAAAGGCCATAGAGATGCTC	542
QY	690	ArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLeuGlyTle	709
Db	543	CGGAGCCGATATCAATGATCTTTCGCTCGAATGACAACAGCCTAGAGTTCTGGGATA	602
QY	710	GlnProThrLeuGlyLyrProProAsnGlnProPro	720
Db	603	CAGCCAACTGAGACCTCTCAACAGCCCTCT	635

RESULT 3
LOCUS BG401683
DEFINITION BG401683
ACCESSION BG401683
VERSION BG401683.1
KEYWORDS EST

879 bp mRNA linear EST 12-MAR-2001
602466350F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4594140 5',
mRNA sequence.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 879)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
<http://image.llnl.gov>
found through the I.M.A.G.E. Consortium/LLNL at:
Plate: L1CM1335 row: d column: 13
High quality sequence stop: 619.
Location/Qualifiers
1..879
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4594140"
/clone_lib="NIH_MGC_75"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Kidney; Vector: pDNR-LIB (Clontech); Site: 1;
SfiI (ggccgcctcgcc); Site: 2: SfiI (ggccattggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGCGCATTTATGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGAGCGCGGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC Library."

BASE COUNT 275 a 177 c 210 g 216 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1.98e-117 Length: 879
Score: 1048.50 Matches: 218
Percent Similarity: 80.50% Conservative: 9
Best Local Similarity: 77.30% Mismatches: 12
Query Match: 27.10% Indels: 45
DB: 12 Gaps: 4

US-09-978-385-2_COPY_19_738 (1-720) x BG401683 (1-879)

QY 370 GlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPheHisGluAlaValGlyGluLe 389
DB 2 CAACCTTTCTGTAGAAATGAGGCTAATGAAAGATTCATGAGCTGTGGGAAATC 61
QY 390 MetSerLeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAsp 409
DB 62 ATGTCACCTTCTGCAGCCACACCACTTAACATTAATCCATGCTGCTGTCACCCGAT 121
QY 410 PheGlnGluAspAsnGluThrGluIleAsnPheLeuLeuLysGlnAlaLeuThrIleVal 429
DB 122 TTTCAAGACACATGAAACAGAAATTAATCACTCTCTCCCAACACACCTCAGATGT- 180
QY 430 GlyThrLeuProPheThrTyrMetLeuGluLysTrpArgTyrPheValPheLysGlyGlu 449
DB 181 GGGACTCTGCATTTACTTACATGTTAAGAAATGAGATGCTCTTTAAAGGGGAA 240
QY 450 IleProLysAspGlnTrpMetLysLysTrpTrpGluMetLysArgGluIleValGlyVal 469
DB 241 ATTCCCAAGACACGATGAGAAAGATGCTGAGCATGAG----- 282
QY 470 ValGluProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSer 489
DB 282 ----- 282

QY 490 AsnAspTyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnGlu 509
DB 283 -----ACCTTTACCAATTCACATTTCCAGAA 309
QY 510 AlaleucysGlnAlaAlaLysHisGlnGlyProLeuHisLysCysAspIleSerAsnSer 529
DB 310 GCACCTTTGTCACGACCTTAACATGAAAGGCGCTCTGCACAAATGATGACATCTCAACCTCT 369
QY 530 ThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgGlyLysSerGluProTyrThr 549
DB 370 ACAGAGCTGAGACAGAAATGTTCAATATGCTGAGGCTTGGAATATGAAACCTGAGAC 429
QY 550 LeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyr 569
DB 430 CTAGCATTTGAAAGATGTTAGACCAAGAACATGATTTAAGCCACTCTCACTAC 489
QY 570 PheGluProLeuPheThrTyrLeuLysAspGlnAsnLysAsnSerPheValGlyTyrSer 589
DB 490 TTTGAGCCCTTATTACCTGCTGCTGAAGACCAAGAACAGAAATTTCTTTGGGATGGAGT 549
QY 590 ThrAspTyrSerProTyrAlaAsp-GlnSerIleLysValArgIleSerLeuLysSerAl 609
DB 550 ACCGACTGGAGTCATATGACAGAACCAAGCATCAAGTGAATGATTAAGCTTAATCATGTC 609
QY 609 aleucGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPheArgSerSerVa 629
DB 610 T---GGAGTATAGCA-TATGACTGAGC---ACATGAATGCTCTCTGCCGATCCTGCT 662
QY 629 lAlaTyr-AlaMetArgLntTyrPheLeuLysValLysAsnGlnMetIleLeuPheGlyG 649
DB 663 GCATATGTCATTTACGAGATGATCTTTAAAGTAAACATGC-----ATGATCTTGGGG 716
QY 649 || 649
DB 717 NG 718

RESULT 4
LOCUS BG962298 766 bp mRNA EST 12-JUN-2001
DEFINITION 602827090P1 NCI_CGAP_CO24 Mus musculus cDNA clone IMAGE:4981606 5',
mRNA sequence.
ACCESSION BG962298
VERSION BG962298.1 GI:14349935
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 766)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM10983 row: d column: 23
High quality sequence stop: 708.
Location/Qualifiers
1..766
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4981606"
/clone_lib="NCI_CGAP_CO24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

FEATURES
source

Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCBI_CGAP Library.*

BASE COUNT 233 a 156 c 179 g 198 t

Alignment Scores:

Pred. No.: 2.2e-115 Length: 766
Score: 1031.00 Matches: 210
Percent Similarity: 88.85% Conservative: 21
Best Local Similarity: 80.77% Mismatches: 23
Query Match: 26.65% Indels: 9
DB: 13 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x BG962298 (1-766)

QY 398 LysHisLeuLysSerIleGlyLeuSerProAspPheGlnGluAspGlnIleuThrGlu 417
DB 2 AACGATCTGAAATCCATGCTCTTCGCAATCCGATTTTCAGAGATGCGAAGACAGAG 61
QY 418 IleAspPheLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrMet 437
DB 62 ATAACTCTCTACTGAAACGCGATGACAAATGTTGGAACACTACCGTTTACTTACATG 121
QY 438 LeuGluLysTrpArgTrpMetValPheLysGlyIleProLysAspGlnTrpMetLys 457
DB 122 TTAGAGAAAGTGGAGTGGATGCTTTCGGGGTGAATTCACAAAGAGAGTGGAGAAA 181
QY 458 LysTrpTrpGlnMetLysArgGluIleValGlyValGluProValProHisAspGlu 477
DB 182 AAGTGTGGAGATGAAAGCGGAGATGTTGGTGTGGAGCTGCTGCTCATGATGAA 241
QY 478 ThrTyrCysAspProAlaSerLeuPheHisValSerAspTrpSerPheIleArgTrp 497
DB 242 ACATACGTCGATCCCTGATCTCTGTCATGTTTAAAGATATACATTCATTCATTCAT 301
QY 498 TyrThrArgThrLeuTrpGlnPheGlnPheGlnGlnAlaLeuGlnAlaAlaLysHis 517
DB 302 TACACAGGACCACTTACCAATTCGAGTTTCAAGAGCTTTGTCAAGGAGTAAAGTAT 361
QY 518 GluGlyProLeuHisLysCysAspIleSerAspSerThrGlnAlaGlnLysLeuPhe 537
DB 362 AATGCTCTCTGCAAAATGTCATCTCAAAATTCACCTAAGCTGGCGAGAGTTGCTC 421
QY 538 AsnMetLeuArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluLysAsnValValGlu 557
DB 422 AAGATGCTGATGCTTGGAAATTCAGAGCTCGACCAAAAGCTTGGAAACATGCTGAGG 481
QY 557 YalaLysAsnMetLysValArgProLeuLeuAsnTrpPheGluProLeuPheThrTrpLe 577
DB 482 AGCAAGGAATATGATATTAACCACTGCTCAATTTCTCCAAACCGTTGTTTGACTGCT 541
QY 577 ULysAspGlnAsnLysAsnSerPhe-ValGly-TrpSerThrAspTrpSerProTyrAla 596
DB 542 GAAAGACAGACAGAAATCTTTCAGTGGGCGGAGAACACTGAAATGAGACCCATATGCC 601
QY 597 AspGlnSerIleLysValArgIle-SerLeuLysSerAlaLeuGlnLysPhe-LysAlaLysArg 616
DB 602 GACCAAGACATTAACAGAGATTAACAGCTGACGCTTGGAGCTTGGAGCTACATGATATG 661
QY 616 LysTrpAsnAsp-AsnGluMetTrpLeuPheArgSerSerValAlaTyrAlaMetArgGln 635
DB 662 ACTGAGACCAACGAAAGATTTCTGCTGCTGATCATCTGT-GCATATGCGATAGAGAC- 719
QY 636 TyrPheLeuLysValLysAsnGlnMetIleLeuPheGlyGluGluAsp 651
DB 720 GTCACTGCTCATATCAAAAACAGACAGTCTTTCTA-GAGAGAGAT 766

RESULT 5 B1913504 763 bp mRNA linear EST 16-Oct-2001
LOCUS B1913504
DEFINITION 603179133P1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5243048 5',
ACCESSION B1913504
mRNA sequence.

VERSION B1913504.1 GI:16177911
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mhc.nci.nih.gov/
AUTHORS 1 (bases 1 to 763)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsb@remail.nih.gov

Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LMNL1613 row: b column: 09
High quality sequence stop: 675.
Location/Qualifiers

FEATURES

source

1..763
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5243048"
/clone_11b="NIH_MGC_121"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb. Insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research genetics tracking code 017. Note: this is a NIH_MGC Library."

BASE COUNT 237 a 167 c 170 g 189 t

Alignment Scores:

Pred. No.: 1.43e-108 Length: 763
Score: 975.50 Matches: 205
Percent Similarity: 90.91% Conservative: 5
Best Local Similarity: 88.74% Mismatches: 12
Query Match: 25.21% Indels: 9
DB: 13 Gaps: 3

US-09-978-385-2_COPY_19_738 (1-720) x B1913504 (1-763)

QY 4 GluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGlnAlaGluAspLeuPheTrp 23
DB 83 GAGGACAGAGCCCAAGACATTTTGGACAAATTTTAAACAGAGCCGAAACCTGTTCTAT 142
QY 24 GlnSerSerLeuAlaSerTrpAsnTrpAsnTrpAsnIleThrGluGluAsnValGlnAsn 43
DB 143 CAAAGTCACTTCTCTTGGATTTTAAACACCAATATTTACGAGAGATTCACAAAAC 202
QY 143 CAAAGTCACTTCTCTTGGATTTTAAACAGAGCTTGGAGCTTGGAGCTTGGAGCT 262
QY 44 MetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAlaGln 63
DB 203 ATGAAATATGCTGGGACAAATGCTGCTTTTAAAGAAAGCTCCACTTGGCCAA 262
QY 64 MetTrpProLeuGlnIleGlnAsnLeuThrValLysGlnGlnLeuGlnAlaLeuGln 83
DB 263 ATGTATCCATTAACAAGAAATTCAGATTCACAGTCAAGCTTCAAGCTTCAAG 322
QY 84 GlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeuAsn 103
DB 323 CAAATGCTCTTCAAGTCTCAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 382
QY 104 ThrMetSerThrIleTrpSerThrGlyLysValCysAsnProAspAsnProGlnGluCys 123

Db 383 ACAATGACACATCTACACTGCGAAAAGTTGTAACCCAGATAATCCACAAGATGC 442
 QY 124 LeuleuLeuGlutProglyLeuAaNgluLeuMetalaAaNsSerLeuApyrAaNgluAa 143
 Db 443 TTAATCTGACACAGAGTTGACTGACATTAATGCAAAAGTTTAACTATGAGAGG 502
 QY 144 LeuTriplaLeu---GluSerTPPArgSerGluValGlyGluLeuAaGrProLeuAa 162
 Db 503 CTCTGGGCTGGCGAAAGCTGGACATCTGAGTGGCGAAGAGCTGAGCCATTAAT 562
 QY 163 GluGluTyrValValLeu-LysAaNgluMetalaAaNgluAaNsHISerGluAaPyrGlu 182
 Db 563 GAGAGATATGTGCTTGAACATGAGATGAGCAAGCAATCATTAATGAGAGATGAG 622
 QY 182 YAspTYrTPArg-GlyAspTYrGluVal-AaNglyVal-AAspGlyTYrAspTYrSerA 201
 Db 623 GGATTTATCGGAGAGAGACTATGAGTAAATGAGGATGAGTATGAGTATGAGCGG 682
 QY 201 gGly-GluLeuIleGlu---AspValGluHISerPhe---GluGluIleLysProLeu 219
 Db 683 CGGACCAAGTTGATTCAGACAGCTGTGGACATACCTTTGACGAGACTTAAACCATAT 742
 QY 219 Yr-GluHISerLeuHISerAa 225
 Db 743 ACTGAACCTCTTCATGCTAT 763

RESULT 6

LOCUS B1561069 676 bp mRNA linear EST 05-SEP-2001
 DEFINITION 603254275F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5295531 5',
 mRNA sequence.
 VERSION B1561069.1 GI:15448383
 KEYWORDS EST.

SOURCE

human.
 ORGANISM

REFERENCE

AUTHORS NIH-MGC http://mhc.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shliraki
 Toshiyuki and Piero Carninci (RIKEN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11750 row: f column: 20
 High quality sequence stop: 674.
 Location/Qualifiers

FEATURES

source

1..676

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5295531"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescript (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtagag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTTNN-3',
 size-selected for average insert size 2.2 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT

215 a 134 c 161 g 166 t

Alignment Scores:

Pred. No.: 2 656-104 Length: 676
 Score: 940.00 Matches: 187
 Percent Similarity: 97.40% Conservative: 0
 Best Local Similarity: 97.40% Mismatches: 1
 Query Match: 24.30% Indels: 4
 DB: 13 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x B1561069 (1-676)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAaPlyrPheAaNsHISerGluValGluAaP 20
 Db 109 TCCACCAATGAGAGAGACAGCCAGACATTTTGGGCAAGTTTAAACCGAAGCCGAGAC 168
 QY 21 LeuPheTYrGlnSerSerLeuAlaSerTPArgSerGluValGlyGluLeuAaGrProLeuAa 40
 Db 169 CTGTCATACAAAGTCACTGCTCTTGGAAATTAACCAATATTAATACGAGAAAT 228
 QY 41 ValGlnAaMetAaNsAlaGlyAspLysTPArgSerAlaPheLeuLysGlu-GlnSerTh 60
 Db 229 GTCCAAACATGATTAATATGCTGGGAGCAATGCTGCTTTTAAAGAGAGAGTCCAC 288
 QY 60 rLeuAlaGlnMetTYrProLeuGlnGluIleGlnAaNsLeuThrValLysLeuGlnLeuG 80
 Db 289 ACTTGCCCAATGTATCCACTACCAAGAAATTCAGAAATCTCAGACTCAAGCTTCAGCTGCA 348
 QY 80 nAlaLeuGlnGlnAaNglySerSerValIleSerGluAaPlyrSerLysAaGrLeuAaNsTh 100
 Db 349 GGCTCTTCAGCAAAATGGCTTTTCAGTCTCTCAGAAAGCAAGCAAGGTTGAAACAC 408
 QY 100 rIleLeuAaNsThrMetSerThrIleTYrSerThGlyValLysAaNsProAaPasnPr 120
 Db 409 AATCTTAATCAATGAGACACCATCTACAGTACGAAAGTTTGTAAACCATATATCC 468
 QY 120 oGluGluCysLeuLeuGluGluProGlyLeuAaNgluIleMetAlaAaNsSerLeuAaPry 140
 Db 469 ACAAGAAATGCTTATCTGAAACAGGTTTGAATGAATATGCAAAACGTTTAACTA 528
 QY 140 rAaNgluAaGrLeuTriplaLeuTriplaGlnSerTPArgSerGluValGlyGluLeuAaGr 160
 Db 529 CAATGAGAGGCTGGGCTTGGGAAAGCTGAGATCTGAGAGTGGCAAGAGCTGAGGCC 588
 QY 160 oLeuTYrGluGluTyrValValLeu-LysAaNgluMetalaAaNgluAaNsHISerGlu 179
 Db 589 ATTATATGAGAGATGCTGCTTGAATAAATGATGAGCAAGCAAAATCATTAATGAGG 648
 QY 180 AspTYr-GlyAspTYrTPArgGlyAsp 188
 Db 649 GACTATGGGGGATTTATGAGAGAGAGC 676

RESULT 7

LOCUS BG722079 635 bp mRNA linear EST 08-MAY-2001
 DEFINITION 602696882F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4830668 5',
 mRNA sequence.
 VERSION BG722079.1 GI:14001266
 KEYWORDS EST.

SOURCE

human.
 ORGANISM

REFERENCE

AUTHORS NIH-MGC http://mhc.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shliraki
 Toshiyuki and Piero Carninci (RIKEN)
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at:

http://image.llnl.gov

Plate: L1M10752 row: c column: 21

High quality sequence stop: 633.

FEATURES

SOURCE

1. 635

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4830668"
/clone_1ib="NIM_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTTTTNN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Gap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIM_MGC Library."

BASE COUNT 202 a 131 c 144 g 158 t

ORIGIN

Alignment Scores:

Pred. No.: 2,2e-102 Length: 635
Score: 924.00 Matches: 180
Percent Similarity: 99.45% Conservative: 0
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 23.88% Indels: 0
DB: 12 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x BG722079 (1-635)

QY 1 SerThrTlleglucIngluAlalysThrPheleuAspLysPheasnhstgluIngluasp 20

Db 94 TCCACCATGTAGAGAACGGCCAAAGACATTTTG-GACAAAGTTTAAACACGAGCGAGAC 152

QY 21 LeuPheTyrgInserSerleuAlaSerTPAsnTyraSnthraSnllthrIngluasn 40

Db 153 CTGTCTATCAAAAGTTTCATCTGCTCTCTGGAATTAATACCCATATTAAGTGAAGAA 212

QY 41 ValGlnAsnMetAsnAlaGlyAspLysTPSerAlaPheleuLysgluInserThr 60

Db 213 GTCCAAACATGAAATATGCTGGGACAAATGCTGCTTTTAAAGAACAGTCCACA 272

QY 61 LeuAlaGlnMetTyrrProleuIngluInleGlnAsnleuThraValLysleuGlnleuGln 80

Db 273 CTTCGCCAAATGATCCACTACCAAGAAATTCAGATCTCACAGTCAGCTTCAGCTGAC 332

QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgleuAsnThr 100

Db 333 GCTCTTCAGCAAAATGGGTCTTTCAGTCTCTCACAAGACAAACGAGTTGACACA 392

QY 101 IleLeuAsnThrMetSerThrIleTyrrSerThrIleLysValLysAsnProAspAsnPro 120

Db 393 ATTCTAATATACAAATGACACATCTACAGTACGAAAGTTTGAACCCAGATTAATCCA 452

QY 121 GlnGlnCysleuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerleuAspTyrr 140

Db 453 CAAAGATGCTTATTAATCTGAACCAAGCTTGAATGAATGAATGAACAGTTTGAAGTAC 512

QY 141 AsnGluArgleuTPAlaTPArgSerTPArgSerGluValGlyLysGlnleuArgPro 160

Db 513 AATGAGAGCGCTGGGCTTGGGAAAGCGAGATCTGAGTGGGCAACGACCTAGAGCCA 572

QY 161 LeuTyrrGlnGluTyrrValIleuLysAsnGlnMetAlaArgAlaAsnHleTyrrGluAsp 180

Db 573 TTATATGGAAGATATGCTGCTTGAATAATGATGACAGCAACATCATTAATGAGAC 632

QY 181 Tyr 181

Db 633 TAT 635

RESULT 8

AA162058

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

The WashU-HMNI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through ILNLT; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:374045
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 446.
Location/Qualifiers

FEATURES

SOURCE

1. 555

/organism="Mus musculus"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:608613"

/clone_1ib="Stratagene mouse skin (#937313)"

/sex="females"

/tissue_type="whole skin"

/dev_stage="11 weeks old"

/lab_host="GOLR (kanamycin resistant)"

/note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI

; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo

dt. Whole skin from 11 week old C57BL/6 female mce.

Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'

adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor

sequence: 5' CTCGAGCTTTTCTTTTCTTTTCTTTT 3'

DB: 9

US-09-978-385-2_COPY_19_738 (1-720) x AA162058 (1-555)

QY 108 IleTyrrSerThrGlyLysValLysAsnProAspAsnProGlnGlnCysleuLeuGln 127

Db 1 ATTATACGATCTGAAAAAGTTTGCAACCCAAAGACCCACAAAGATCTTATTAATCTTGAC 60

QY 128 ProGlyLeuAsnGlnIleMetAlaAsnSerleuAspTyrrAsnGluArgleuTPAlaTP 147

Db 61 CCAGGATTTGATGAAATATGATGGCAGCAACACA-GACTACAAAGCTAGGCTCTGGCATGG 119

QY 148 GluSerTPArgSerGluValGlyLysGlnleuArgProleuTyrrGluGluTyrrValVal 167

Db 120 GAGGGCTGGAGGGCTGGTGGCAAGCAGCTGAGGCGCTGTATGAAAGATGTGGTC 179
Qy 168 LeuLysAngLumetaLaragaLAsnHsTYRGLUAspTYRGLYAspTYRTPArgLY 187
Db 180 CTGAAGAACAGATGCGCAGAGCAACCAATTATACGACTATGGATTTATGGAGAGG 239
Qy 188 AspTYRGLUValAsnGLYValAspGLYTYRAspTYRSerArgGLYGLNLeuILEGLUAsp 207
Db 240 GACTATGACAGCAGAGGAGCAGTGGCTCAACTATACCCGTAACCAAGTATGACAT 299
Qy 208 ValGLUHisThrPheGLUValLeuLYSProLeuTYRGLUHisLeuHsAlaTYRValArg 227
Db 300 GTGAAACGTAACCTCTCCAGAGATCAAGCATTGTATGAGCACTTCAAGCTATGTGAGG 359
Qy 228 AlaLYSLeuMetAsnAlaTYRProSerTYRILESerProILEGLYCYLeuProAlaHis 247
Db 360 AGGAAAGTGTATGATGATACCTACCTCTCTCAATACACACCCCACTGATGCTCCCTCCCAT 419
Qy 248 LeuLeuGLYAspMetLeuTYRArgPheThrPheAsnLeuTYRSerLeuThrValProPhe 267
Db 420 TTGCTGTGTGATATGTGGGTGATGTTGGCAAAATCTGTACCTTGTACTGTTCCCTTT 479
Qy 268 GLYGLNLYSProAsnILEAspValThrAspAlaMetValAspGLNAlaThrAspAlaGLN 287
Db 480 GCACAGAAACCAACATAGATGTTCTACTGATGATGATGATGATGATGATGATGATGATGAT 539
Qy 288 ArgILEPheLYSGlu 292
Db 540 AGGATATGTCAAGAG 554
RESULT 9
AK008530
LOCUS
DEFINITION Mus musculus adult male small intestine cDNA, RIKEN full-length enriched library, clone:2010305105;homolog to ANGIOGENIN CONVERTING ENZYME-LIKE PROTEIN (ACE-RELATED CARBOXYPEPTIDASE ACE2), full insert sequence.
ACCESSION AK008530
VERSION AK008530.1 GI:12842766
KEYWORDS HTC, CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) adult male small intestine cDNA to clone:2010305105.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, N., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
AUTHORS Shibata, K., Itoh, N., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasai, Y., Fujisake, S., Inoue, K., Togawa, T., Izawa, M., Ohara, E., Watanabe, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawal, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913

PUBMED 11076861
REFERENCE
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, O., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Goto, T., Bono, H., Kasukawa, T., Saito, R., Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Riechmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nakai, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stabili, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Holtmann, M., Hume, D. A., Kamuya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
REFERENCE
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, K., Harada, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, K. I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
FEATURES
source location/Qualifiers
 1..1603
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:2010305105"
 /db_xref="MGI:1902242"
 /db_xref="taxon:10090"
 /clone="2010305105"
 /sex="male"
 /issue_type="small intestine"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 49..798
 /note="data source:SPR, source key:09NRA7, evidence:ISS homolog to ANGIOGENIN CONVERTING ENZYME-LIKE PROTEIN (ACE-RELATED CARBOXYPEPTIDASE ACE2)
 putative"
 /codon_start=1
 /protein_id="BAB25723.1"
 /db_xref="GI:12842766"
 /db_xref="MGI:1917258"
 /translation="MLSGNSEPWTGLNENYVGRNDVPLNTPOPEFDWKEQNR NFGVWNTNEMSPYADSIKVAISLKALGANAYEWNTNEMFLFRSSVAYAMRYFSII

KN0VPLEEDVRSVLDKRPSEFEVTSPOVSDVDPREVEDALRMSGRINDVFG
 LINDSEIFGHPLEPPYOPVPTIMLIFGVVALVYVGLILLVTGKGRKKNET
 KRENSYDSMDIGKGSNAGFONSDDAQS¹⁵⁸⁴

polyA_signal
 1584..1589
 /note="putative"
 1603
 polyA_site
 /note="putative"
 1603

BASE COUNT 465 a 321 c 366 g 450 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 4.95e-90 Length: 1603
 Score: 829.00 Matches: 155
 Percent Similarity: 88.83% Conservative: 20
 Best Local Similarity: 78.68% Mismatches: 22
 Query Match: 21.43% Indels: 0
 DB: 11 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AK008530 (1-1603)

QY 524 CysAspIleSerAnserrThrgluAlaGlyIleLeuPheAsnMetLeuArgLeuGly 543
 Db 4 TGTGACATCTCAAAATTCAGTGAAGCTGGGCAAGAGTTGCTCAAGATGCTGAATCTTGA 63
 QY 544 LysSerGluProThrIleuAlaLeuGluAsnValAlaGlyAlaLysAsnMetAsnVal 563
 Db 64 AATTCAAGCCCTGGACCAAGCCTTGAAAGTGGTAGCAGCAAGAGATATGATGTA 123
 QY 564 ArgProLeuLeuAsnTrpPheGluProLeuPheThrIleuLysAsnGlnAsnLysAsn 583
 Db 124 AATACACCTGCTCAATTCCTTCAACCTGTTGACCTGGGCAAGAGCAAGCAAGT 183
 QY 584 SerPheValGlyTrpSerThrAspTrpSerProTrpAlaAspIleSerIleLysValArg 603
 Db 184 TCTTTTGTGGGTGGAACCTGATGAGAGCCCATATGCCACCAACCAATTAAGTGAGG 243
 QY 604 IleSerLeuLysSerAlaLeuGlyAspLysAlaTrpGluTrpAsnAspAsnGluMetTrp 623
 Db 244 ATAAGCCTAAATAGCTCTTGAGCTAAATGCAATGAAATGAGCAACCAAGAAATGTTTC 303
 QY 624 LeuPheArgSerSerValAlaTrpAlaMetArgGlnTrpPheLeuLysValLysAsnGln 643
 Db 304 CTGTCCGATCATCTGTGATGATGCAATGAGAAAGTATTTTCAATAATCAAAAACAG 363
 QY 644 MetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPhe 663
 Db 364 ACAGTCTCTTTCTTCAAGAGAGAGATGACGAGTGAATTTGAAACCAAGAGTCTCTTC 423
 QY 664 AsnPhePheValThrAlaProLysAsnValSerSerIleIleProArgTrpGluValGlu 683
 Db 424 TACTTCTTTGTACCTCACCCCAAAATGTGTCTGATGATGCTAGAGAGTGAAGTGA 483
 QY 684 LysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSer 703
 Db 484 GATGCCATCAGAGATCTCGGGGCGGCATCATATGATGCTTTGGCCTGAATGATTAACGC 543
 QY 704 LeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 544 CTGAGGTTTCTGGGATTCACCCACACTGAGCAGCACTTACAGCCTCCT 594

RESULT 10
 B1561359 609 bp mRNA linear EST 05-SEP-2001
 LOCUS B1561359
 DEFINITION 60325121F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5297380 5',
 mRNA sequence.
 ACCESSION B1561359
 VERSION B1561359.1 GI:15448673
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 609)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cegaps-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshitsuyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LAM11752 row: 1 column: 05
 High quality sequence stop: 606.
 Location/Qualifiers
 1..609
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5297380"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gcgag
) Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.2 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 193 a 128 c 137 g 151 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3.89e-90 Length: 609
 Score: 824.00 Matches: 169
 Percent Similarity: 97.70% Conservative: 1
 Best Local Similarity: 97.13% Mismatches: 1
 Query Match: 21.30% Indels: 4
 DB: 13 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x B1561359 (1-609)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 Db 94 TCCACCATTTGAGGAACAGGCAAGACATTTGTGACAAAGTTTAAACCAAGAGCGAAGAC 153
 QY 21 LeuPheTrpGlnSerSerIleuAlaSerTrpAsnTrpAsnThrAsnIleThrGluGluAsn 40
 Db 154 CTGTCTATCAAAAGTTCACCTTGTCTGTGAATTAATTAACCAATTAATGTAAGAGAT 213
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerTh 60
 Db 214 GTCCAAAACATGATATATGCTGGGGGACAAATGTCTGCTTTTAAAGAGAAAGTCCAC 273
 QY 60 rIleuAlaGluMetTrpProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuG 80
 Db 274 ACTTGCACCAATGATACCTACAGAAATTCGAATCTACAGTCAAGCTTAAGCTGCA 333
 QY 80 nAlaLeuGlnGln-AsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnTr 100
 Db 334 GGCTCTTACGACCAAAAGGGTCTTCAGTGTCTCAAGAAAGCAAGAGCAAGGTTGAACA 393
 QY 100 hr-IleuAsnThrMetSerThrIleTrpSerThrGlyLysValLysAsnProAspAsn 119
 Db 394 CGAATTCATAATTAAGAGACACATCTAAGTACAGTGAAGAGTTGTAACCCAGATAT 453
 QY 120 ProGlnGluCysLeuLeuGluGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAsp 139
 Db 454 CCACAGAGATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 513
 QY 140 TyraAsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArg 159

```

Db      514 TACATGAGAGAGCTGCTGGCTGGAGAAAGCTGAGATCTGAGTGGAG-CAAGTGAAG 572
Qy      160 ProleutyrglulgluYrValleuLysAsnGlu 171
Db      573 CCATTATATGAGAGATGTGTGCTTTGAAAAATGAG 608

RESULT 11
LOCUS   BF789159                      965 bp    mRNA    linear    EST 12-JAN-2001
DEFINITION 602105037f1 NCL CGAP_K1d14 Mus musculus cDNA clone IMAGE:4223312
VERSION   BF789159
ACCESSION BF789159.1 GI:12094195
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 965)
AUTHORS   NIH-MGC http://mgc.ncl.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: jeffrey E. Green, M.D.
          CDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          http://image.llnl.gov
          Plate: LLM9811 row: 1 column: 09
          High quality sequence stop: 676.
          Location/Qualifiers
            1..965
              /organism="Mus musculus"
              /strain="FVB/N"
              /db_xref="taxon:10090"
              /clone="IMAGE:4223312"
              /clone_1lb="NCL CGAP_K1d14"
              /lab_host="DH10B (TI phage-resistant)"
              /note="Organ: Kidney; Vector: pCMV-SPORT6; Site:1; NotI;
              Site_2: SalI; Cloned unidirectionally. Primer: 011go dt.
              Average insert size 1.75 kb. Constructed by Life
              Technologies. Note: this is a NCL CGAP library."
BASE COUNT 286 a 213 c 253 g 212 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1.92e-88      Length: 965
Score: 813.00           Matches: 166
Percent Similarity: 79.82% Conservative: 16
Best Local Similarity: 72.81% Mismatches: 43
Query Match: 21.01%      Indels: 4
DB: 12                  Gaps: 2

US-09-978-385-2_COPY_19_738 (1-720) x BF789159 (1-965)
Qy      1 SerThrIleGluGluGlnAlaIleValSerThrPheLeuAspLysPheAsnHisGluAlaGluASP 20
Db      74 TCCCTCACCAGGAGAAATGCCAAGACATTTTAAACAATTTATCATCGAAGCTGAAGAC 133
Qy      21 LeuPheTyrgInSerSerIleuAlaSerTrpAsnTyrsAnthrAsnIleThrGluGluAsn 40
Db      134 CTGTCTATCAAGATGCTCTGCTTGGATTATATATCTAACATCTACTGAAGAAAT 193
Qy      41 ValGlnAsnMetAnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
Db      194 GCCCAAAAGATGATGAGCTGCAGCCCAATGGCTGCTTATGAGAAACAGCTTAAG 253
Qy      61 LeuAlaGlnMetTyrgProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80

```

```

Db      254 ACTGCCAAAGTTTCTCATTACAGAAATCCAGACTCCGATCATCAAGCTCAACTACAG 313
Qy      81 AlaleuGlnLinsnglySerValleuSerGluLysAspLysArgLeuAsnThr 100
Db      314 GCCCTTCAGCAAGATGGGTCTTCAGCATCTCAGCAGACACAAACAAACAGTTGACACA 373
Qy      101 IleLeuAsnThrMetSerThrIleTySerThrGlyLysValCysAsnProAspAsnPro 120
Db      374 ATTCTGAACACCAATGAGACCACTTACACTGGAAGATTTCACAAACCAAGACCA 433
Qy      121 GlnGluCysLeuLeuLeuGluGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTy 140
Db      434 CAAGAATGCTTATTATTACATTCAGCCAGATTTGATGAATTAATGACACAAAGACAGACTAC 493
Qy      141 AsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
Db      494 AACTCTAGAGCTCGGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 553
Qy      161 LeuTyrgLulgluYrValleuLysAsnGluMetAla-ArgAlaAsnHisTyrgLus 180
Db      554 TTGTATGAGAGAGATGTGTCTCTGAAAAACGAGATGCGAAAGACAAACAAATTTAACA 613
Qy      180 PTrGlyAspTyrgTrpArgGlyAspTyrgLulgluValAsnGlyValAspGlyTyrgAspTy 200
Db      614 CTATGGGAGATTATTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 672
Qy      200 rArgGlyLinsnglyLeuGluLysPheValGluHisThrPheGluGluIleLysProLeuTyrg 220
Db      673 CCCGTACAGCTGATTTGAGA---TGGAACGTACTTTCGCAAAA---TCACGATTTGGTGA 726
Qy      220 whLseuHisAlaTyrgValarg 227
Db      727 GCCTCTTCATGCGCTATGTGAGG 748

RESULT 12
LOCUS   BG428060                      452 bp    mRNA    linear    EST 14-MAR-2001
DEFINITION 602501471f1 NIH-MGC_75 Homo sapiens cDNA clone IMAGE:4615121 5'
VERSION   BG428060
ACCESSION BG428060.1 GI:13334566
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 452)
AUTHORS   NIH-MGC http://mgc.ncl.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: CLONTECH Laboratories, Inc.
          CDNA Library Preparation: CLONTECH Laboratories, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L1CML167 row: n column: 18
          High quality sequence stop: 450.
          Location/Qualifiers
            1..452
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:4615121"
              /clone_1lb="NIH-MGC_75"
              /lab_host="DH10B (TI phage-resistant)"
              /note="Organ: Kidney; Vector: pNR-LIB (Clontech); Site:1;
              Site1 (ggccgcctggcgc); Site_2: SfiI (ggccatcatggcc); 5' and
              3' adaptors were used in cloning as follows: 5' adaptor
              sequence: 5'-CACGCCCATTTATGAGCC-3' and 3' adaptor sequence:
              5'-ATTCTAGAGCGCAGGCGGCGCAGCATG-dT(30)BN-3' (where B = A,

```

C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library.

BASE COUNT 144 a 81 c 103 g 124 t

ORIGIN

Alignment Scores:

Pred. No.: 4.45e-85 Length: 452
Score: 781.00 Matches: 150
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.19% Indels: 0
Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x BG428060 (1-452)

QY 552 LeuGluAspValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGlu 571
DB 2 TTGGAAAGAGTTTATGAGCAAGAAACATGATGAGCCACTGCTCAACTTGG 61
QY 572 ProLeuPheThrTyrLeuLysAspGlnAsnLysAsnSerPheValGlyTyrSerThrAsp 591
DB 62 CCTTATTTTACCTGGCTGAAAGACACCAAGAAATTTCTTGTGGATGAGTACCGAC 121
QY 592 TTPSerProTyrAlaAspGlnSerIleLysValArgIleSerLeuLysSerAlaLeuGly 611
DB 122 TGGAGTCCATATCAGACCAAGCATCAAGTATGAGTAAAGCTTAATAGCTTGG 181
QY 612 AspLysAlaTyrGluTyrPheAsnAspAsnGluMetTyrLeuPheArgSerValAlaTyr 631
DB 182 GATTAAGCATATGATGAGCAAGCAAGAAATGATGATGCTGCTGATCATCTGTCATAT 241
QY 632 AlaMetArgGlnTyrPheLeuLysValLysAsnGluMetIleLeuPheGlyGluGluAsp 651
DB 242 GCTATGAGGAGCATCTTTTAAAGTAAATCAGATGATCTTTTGGGAGAGGAT 301
QY 652 ValArgValAlaAsnLeuLysProArgIleSerPheAsnSerPheValThrAlaProLys 671
DB 302 GTGGAGATGGCTAATTTGAAACCAAGATCTCTTATATTTCTTGTGCACTGACCTGAA 361
QY 672 AsnValSerAspIleIleProArgTyrGluValGlyLysAlaIleArgMetSerArgSer 691
DB 362 AATGTGCTGATATCATCTCTAGAACTGAAGTTGAAAGCCATCAGATGTCCTGGAGC 421
QY 692 ArgIleAsnAspAlaPheArgLeuAsnAsp 701
DB 422 CGTATCAATGATGCTTCCGTCTGATGAC 451

RESULT 13

BM030353 556 bp mRNA linear EST 05-NOV-2001

LOCUS 488973 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.

DEFINITION BM030353

ACCESSION BM030353.1 GI:16743923

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 556)

Smith, T.P.L., Grose, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,

Caas, E., Wray, J.E., White, J., Cho, J., Fahnenkrug, S.C., Bennett,

G.L., Heaton, M.P., Isagrd, W.W., Rohrer, G.A., Chitto-McCown, C.G.,

Peterson, G., Holt, I., Karameycheva, S., Liang, F., Quackenbush, J. and

Keele, J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA

libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

21180013

COMMENT

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemil@marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
and -minmatch 12 options.

PCR primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCAGTCCAGCAGC

Plate: 120 Row: B Column: 16

Seq primer: ATTATGATGACATATG.

Location/Qualifiers

1..556

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 2BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;

Library made from pooled tissue from testis, thymus,

semilendonsus muscle, longissimus muscle, pancreas,

adrenal, and endometrium."

DB: 13

Indels: 1

Gaps: 1

US-09-978-385-2_COPY_19_738 (1-720) x BM030353 (1-556)

QY 53 AlaPheLeuLysGluGlnSerThrLeuAlaGlnMetTyrProLeuGlnGluLeuGlnAsn 72
DB 2 GCTTTTATGAGAAAGATGCTCCGATGCGCAAACTTACTGCTGAAAGAAATTTAGAT 61
QY 73 LeuThrValLysLeuGlnGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGlu 92
DB 62 CTCACACTCAAGGCTAATTTGAAGGCTTTCAGCATGAGGAGGACCTCAGCTCGGCA 121
QY 93 AspLysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGly 112
DB 122 GAGAAAGCAAAAGATTTGAACACGATTTTAAATTAATGAGCAACCATTTACAGTCTGG 181
QY 113 LysValCysAsnProAspAsnProGlnGluCysLeuLeuLeuGlnProGlyLeuAsnGlu 132
DB 182 AAGTTTGGACCA--AATACACAGAGAGTGTGCTTGCACCTTGAAACAGTTTATGATGC 238
QY 133 IleMetAlaAsnSerLeuAspTyrAsnGluTyrAlaLeuPheLeuPheLeuPheArgSer 152
DB 239 ATATGGAAGAAACAGACAGACGACTAATCGAGGCTCTGGGCTTGGGAGGCTGAGGCT 298
QY 153 GluValGlyLysGlnLeuArgProLeuTyrGluGlnTyrValValLeuLysAsnGluMet 172
DB 299 GAAATTTGGCAAGCAACCAAGCCATTTGAAGAGATGCTGCTTGAAGAAATGAGATG 358
QY 173 AlaArgAlaAsnHisTyrGluAspTyrGlyAspTyrThrArgGlyLysPyrGluValAsn 192
DB 359 GCAAGACCCACAACTTTATGAGACTACGAGGAGGACTATTGAGAGGATTTATGAGTACT 418
QY 193 G1yValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGluHisThrPhe 212
DB 419 GGGGAGGAGGACTATGACTACAGCCGAGCAGTGTGTAAGATGATGAGACACGACCTTT 478
QY 213 GluGluLeuLysProLeuTyrGluHisLeuHisAlaTyrValArgAlaLysMetAsn 232
DB 479 GCAGAGATTAAACCATTTATGAAACAACCTTCACTTATGTGAGGGCAAGCTTGATGCAAT 538

QY 233 AlatyProserTyrille 238
 DB 539 ACCATCCCTCTATATC 556
 RESULT 14
 BB652968
 LOCUS
 DEFINITION BB652968 686 bp mRNA linear EST 31-AUG-2001
 BB652968 RIKEN full-length enriched, adult male hippocampus Mus
 musculus cDNA clone C630041D11 5', mRNA sequence.
 BB652968
 VERSION BB652968.1 GI:15402926
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 686)
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
 Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Koda
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki
 D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyai, T.,
 Muramatsu, M., and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 Unpublished (2001)
 TITILE Contact: Yoshihide Hayashizaki
 JOURNAL Laboratory for Genome Exploration Research Group, RIKEN genomic
 COMMENT Science Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multiplexed sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A., Aizawa
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T.,
 Ishii, Y., and Hayashizaki, Y.
 Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
 Func. Genomics 2 pre, L72-L86 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 e mouse tissues.
 FEATURES
 SOURCE Location/Qualifiers
 1. 686
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="C630041D11"
 /clone_id="RIKEN full-length enriched, adult male
 hippocampus"
 /sex="male"
 /tissue_type="hippocampus"
 /dev_stage="adult"
 /lab_host="SOLR"
 /note="Site_1: XhoI; Site_2: BamHI; cDNA library was

BASE COUNT 211 a 151 c 169 g 151 t 4 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,41e-80
 Score: 741.00
 Percent Similarity: 80.30%
 Best Local Similarity: 73.23%
 Query Match: 19.15%
 DB: 10 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x BB652968 (1-686)
 QY 1 SerhrllgllugllugllalalysThrPheLeuaspLysPheasnHsglualagLasp 20
 DB 86 TCCTCCACCGGGAATTCAGAGACATTTTAAACCACTTATCAGAGAGCTGAAGC 145
 QY 21 LeuphetYrgInserSerleuAlaSerTyrAsnThrAsnThrlnlePhrgluuan 40
 DB 146 CTGCTTATCAAGATTCACCTGCTTCTTGGAATTTATATCTATCACTGAGAAAT 205
 QY 41 ValGlnasnMetasnAlaGlyaspLysTyrSerAlaPheLeuaspLysInserThr 60
 DB 206 GCCCAAAAGATGAGTGAGCGCTGCAGCCAAATGCTGCTTATGAGACAGCTGTAAG 265
 QY 61 LeuAlaglnMetYrProLeuglnGlnGlnGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 266 ACAGCCCAAGATTCACCTACACAGAAATCCAGACTCCATGCTCAAGGCTCAACAGC 325
 QY 81 AlaleuGlnGlnasnGlySerSerValLeuSerGluaspLysSerLysArgLeuasnThr 100
 DB 326 GCCCTTCAGCAAGATGCTGCTTCTGACGACCTCTGACGACAGACAGACAAAGTTGAAC 385
 QY 101 IleLeuasnThrMetSerThrIleTyrSerThrGlyLysValCysasnProaspasnPro 120
 DB 386 ATTCTGACACACATGAGACACCATTTACATCTGCAAAAGTTGCAACCCAAAGAACCA 445
 QY 121 GlnGluCysLeuLeuLeuGlnLuproGlyLeuAsnGlnIleMetAlaAsnSerLeuaspThr 140
 DB 446 CAAGAAATGCTTATTTACTTGGCAGAGATTTGATGAATTAATGAGCAAGACAGACTAC 505
 QY 141 AsnGluArgLeuThrAlaTrpGlnSerTyrParGserGluValGlyLysGlnLeuArgPro 160
 DB 506 AACTTAGGCTCTGGGCGATGAGAGGCGCTGAGCGCTGAGCGCAACAGCTGAGGGCC 565
 QY 161 LeuTyrGluGluTyrValValLeuLysasnGlnLueAlaArgAlaAsnHsTyrGluAs 180
 DB 566 GTGTATGAGAGATTAAGGGGCGCTGAAAAACAGAGAGGCAAGAACCAATTTATTAACA 625
 QY 180 pTyrGlyaspTyrTyrArgGlyaspTyrGluValasnglyValaspGly 196
 DB 626 CTATGGGATTTATTTGANNAGGGGACTTTGAAGCCNAGGGAGCAAGG 675
 RESULT 15
 BIR26471
 LOCUS
 DEFINITION BIR26471 800 bp mRNA linear EST 04-OCT-2001
 60307617791 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5167987 5',
 mRNA sequence.
 ACCESSION BIR26471
 VERSION BIR26471.1 GI:15938021
 KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 800)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1417 row: 3 column: 20
High quality sequence stop: 797.
Location/Qualifiers
1. 800
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5167987"
/clone_lib="NIH-MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: This is a NIH-MGC library."
BASE COUNT 169 a 244 c 230 g 157 t
ORIGIN
Alignment Scores:
Pred. No.: 3,58e-73 Length: 800
Score: 687.50 Matches: 128
Percent Similarity: 68.30% Conservative: 53
Best Local Similarity: 48.30% Mismatches: 81
Query Match: 17.7% Indels: 3
DB: 13 Gaps: 3
US-09-978-385-2_COPY_19_738 (1-720) x B1826471 (1-800)
QY 212 PheGluGluLeuLysProLeuTYRGIuHisLeuHisAlaTYRValArgAlaLysLeuMet 231
Db 9 TTCAGAGGCTGCAGCCACTTACCTACCTGATGCTGACCGCCGCGCTGCAC 68
QY 232 AsnAlaLysr--ProSerTYRLeuSerProIlleGlyCysLeuProAlaHisLeuLeuGly 250
Db 69 CGTACTACAGGGGCGCCGACATCACTGAGGGGCGCCATCTCTGCTCACTGCTGGG 128
QY 251 AspMetTrpGlyArgPheTrpHisLeuTYRLeuValProPheGlyGlnLys 270
Db 129 AACATGTGGCGGACCTGCTCAACATCATGACTTGGTGGGCGCTTCCCTTCAGCC 188
QY 271 ProAsnLeuAspValThrAspAlaMetValAspGlnAlaTrpAspAlaGlnArgIlePhe 290
Db 189 CCTCTGATGACACCAAGAGGCTATGCTAAAGCAGGGCTGAGCGCCAGAGAGATGTT 248
QY 291 LysGluLacIuLysPhePheValSerValGlyLeuProAsnMetTrpGlnGlyPheTrp 310
Db 249 AAGAGGCTGATGATTTCTTCACTCCCTGGGCTGTGGCCGCTCGTCAAGTTCTGG 308
QY 311 GluAsnSerMetLeuTrpAspProGlyAsnValGlnLysAlaValCysHisProThrAla 330
Db 309 AACAGTGTATGCTGGAGAGCAACGACGCGGGAGAGTGTCTGCCACGCTCGGCC 368

QY 331 TrpAspLeuGlyLysGly---AspPheArgIleLeuMetCysThrLysValThrMetAsp 349
Db 369 TGGAGCTTTCACAAAGAGAGCTCCGATCAAGACATGCAACCGACCTGAGTGGAG 428
QY 350 AspPheLeuThrAlaHisHisGluMetGlyHisIleGlnTYRAspMetAlaTYRAlaAla 369
Db 429 GACCTGTGTGGGCCACACGAAATGGGCCACATTCAGTATTCATGACGATCAAGAAC 488
QY 370 GlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPheHisGluAlaValGlyGluIle 389
Db 489 TTACTGTGGCTTGGAGGAGGAGGTCGCAACCCCGCTTCCATGTAGGCGCATTTGGAGCGTG 548
QY 390 MetSerLeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAsp 409
Db 549 GTAGCCCTGTAGTGTCTACGCCCAAGCAGCTGACATCTCAACCTGCTGAGCAGTGAG 608
QY 410 PheGlnGluAspAsnGluThrGluLeuAsnProPheLeuLeuLysGlnAlaLeuThrIleVal 429
Db 609 GTGGCAGCGAC---GAGCATGACATCAACTTTCGATGAGAGATGGCCCTTGACAAAGATC 665
QY 430 GlyThrLeuProPheTrpTYRMetLeuGluLysTrpArgTrpMetValPheGlyGlu 449
Db 666 GCCTTATCCCTTCAGCTACCTGCTGATCAGTGGCGCTGAGGATTTGAT-GGAAGC 724
QY 450 IleProLysAspGlnTrpMetLysLysTrpTrpGluMetLysArgGluIleValGlyVal 469
Db 725 ATCCACCAAGGAGACTATACCGAGAGTGTGAGCTCAGGCTGAGGCTGACAGGAGCT 784
QY 470 ValGluProValPro 474
Db 785 CTGCCCCAGTGCCTC 799
RESULT 16
B1415293 900 bp mRNA linear EST 14-AUG-2001
LOCUS 60287289F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5143199 5'
DEFINITION mRNA sequence.
ACCESSION B1415293
VERSION B1415293.1 GI:15176216
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 900)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1353 row: a column: 24
High quality sequence start: 9
High quality sequence stop: 873.
Location/Qualifiers
1. 900
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5143199"
/clone_lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pTZ19.3d-Pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRV; 1st strand cDNA was prepared from mRNA obtained from pooled

Email: cgapds-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LLM11643 row: f column: 13
High quality sequence start: 37
High quality sequence stop: 848.
Location/Qualifiers

```

FEATURES
    source
        1..848
            /organism="Mus musculus"
            /strain="CZECH II"
            /db_xref="taxon:10090"
            /clone="IMAGE:5254668"
            /clone_1ib="NCI CGAP Lu33"
            /tissue_type="Pooled lung tumors"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: Lung; Vector: p773D-pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5', ngtttccaactctgaagggaggccgcctcgattggtttttttttt 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT      213 a      228 c      228 g      228 t

ORIGIN
Alignment Scores:
Pred. No.:          6.16e-70           Length:         848
Score:              661.50             Matches:         127
Percent Similarity: 64.10%             Conservative:    48
Best Local Similarity: 46.52%           Mismatches:     87
Query Match:        17.10%             Indels:         11
DB:                 13                  Gaps:           5

US-09-978-385-2.COPY.19_738 (1-720) x BI904144 (1-848)
OY       319   G|YASnvaLGlnvsaLa-----ValCySHISProThraLaTPAsp 332
               |||||.....:|||                |||||.....:|||
Db       32   GGTATGTTTAGAACAACCCACCAGTGAAGGAGGAGTGCGTCACATCCCTCAGCGTGGAGC 91
OY       333   Leugllylsely---AsphearglyleUwMetCysThrlysValTLTMtaPaaspe 351
               |||                    |||
Db       92   TTTCACAAcGGCAAGACTTCAGGATCAACAGCATACCCTCTGTGAACATGGAAGACTT 151
OY       352   LeuthrAlaHshISglumetGLyHIsileGInTyArPMetaLaTyraLaalaginPro 371
               ::|  |||||.....:|||                |||
Db       152  GTGATAGCGCACCAAGAANAATGGGCCACATCCAGTATTTCATGACAGTCAAAAGACTTCCC 211
OY       372   PheLeuLeuArGsnglYalaasngluGlpHehISgluaLaYaalgJLuIlleMerSer 391
               |||  |||||.....:|||                |||||.....:|||
Db       212  GTGACTTTCCGGGAGGGTGCACACCTGTGGTTTCATGAAGCATATG--GAGATTAAGCT 268
OY       392   LeuSerAlaLaThrProLySHISleuLySserillegLYLeuLeuSerProaspPheGln 411
               |||||  ::|||.....:|||                |||||.....:|||
Db       269  CTCCTCAGTGCTACCCCCCAAGCATCATATACAGTCTCAACCTGCTTACAGCATGAG--GGC 325
OY       412   GluAspaNsnglurHguIlleaSpheLeuLeuLySGlnAlaLeuThrIlleValGlYthr 431
               |||  ::|||.....:|||                ::|
Db       326  AGTGCGTCAcGGTATGATCAACACTTTCATATAGATAGATGGCCCTGACAAAGATGCGCTTT 386
OY       432   LeuProPheHrrTyMtleuGLuLySTPrArgTrpMeValPheLySGlyGluIlePro 451
               ::|||.....:|||                |||||.....:|||
Db       386  ATCCCCCTTACACTTCTCATCGACGACAGTGGCGGTGTGATGGAAAGATGACC 445
OY       452   lysAspGlnTrpMetLysLySTTrpGluMetLysArgGluIleValGlyValValGlu 471

```

Alignment Scores:			
Pred. No.:	1,93e-68	Length:	741
Score:	648.50	Matches:	151
Percent Similarity:	86.67%	Conservative:	5
Best Local Similarity:	83.89%	Mismatches:	11
Query Match:	16.76%	Indels:	13
DB:	12	Gaps:	4
US-09-978-385-2_COPY19_738 (1-720) x BG772384 (1-741)			
QY	1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHISgluAlaGluAsp		20
DB	112 TCACACCAATGAGGAACAGGCCAACAGACTTTTGTGACAACTTAAACACGAGAGCCAAAC		177
QY	21 LeuPheTYrGlnSerSerLeuAlaSerThrAspTYrAsnIleThGluAsn		40
DB	172 CGTTCTCATCAAACTTCACTGCTCTTGGAAATTTAAACCAAAATTTAGTAAGAGAAAT		231
QY	41 ValGlnsMetAsnAsnAlaGlyAspLysTYrSPSerAlaPheLeuLysGluGlnSerThr		60
DB	232 GTCCAAAACATGAATATATGCTGGGGCAAAATGCTGCCTTTATACGGAACACTCCACA		291
QY	61 LeuAlaGln-MetTYrProLeuGlnIleGlnAsnLeuThrValLysLeuGlnLeuG		80
DB	292 CTTCGCCCAATATGATTCACCTACACAAATTCAGAAATTCACAGTCAAGCTTCAAGTCGA		351
QY	80 nAlaLeuGlnGlnAsn---GlySerSerValLeuSerGluAspLysSerLys-ArgLeu		99
DB	352 GGCCTTCACACACACACTGGGTCTTCAGTGTCTCAACAAACAAAGACCAACAGCTTGA		411
QY	99 snThrIleLeu-LeuAsnThr-MetSerThrIleTYrSerThrGly-LysValCysAsnPro		118
DB	412 ACACAAATTCCAATACACATACACAGACCCACTTACAGTACTGGAACCAACAGTCTGAACCCAG		471
QY	118 sPAsnProGlnGlnCysLeuLeuLeuGluProGlyLeuAsnGlu-IleMetAlaAsnSer		137
DB	472 ATATATCCAAAGAAATCTTATTACTTGAACCAAGGTTTAAATGACACTACCTGGCAACAGT		531
QY	138 LeuAspTYrAsn---GluArgLeuThrPheAlaTrpGluSer-TTPrArgSerGluValGly		156
DB	532 TTGACACTACCACTGACAGAGGCTCTGGGGCTGGGGAAGACCTGGAGAACT---GACAGTGG		588
QY	156 sGlnLeuArg-----ProLeuTYrGlu-GluTYrValValLeu		168
DB	589 TGCAGCACGCTGAGGCCACTTATATGAGAGAGTATATGAGGGGCTT 632		
RESULT 20	848 bp	MRNA	linear
B1766340			EST 25-SEP-2001
LOCUS	603052244F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5202083 5',		
DEFINITION	603052244F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5202083 5',		
ACCESSION	B1766340		
VERSION	B1766340.1		
KEYWORDS	GI:15757918		
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	1 (bases 1 to 848)		
COMMENT	NIH-MGC http://mgs.nci.nih.gov/.		
	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cga@b-remail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLAM11506		
	Row: 9		
	Column: 12		

